



## Essays in Microeconomics

Lorenzo Verstraeten

Thesis submitted for assessment with a view to obtaining the degree of  
Doctor of Economics of the European University Institute

Florence, 06 May 2019



European University Institute  
**Department of Economics**

## Essays in Microeconomics

Lorenzo Verstraeten

Thesis submitted for assessment with a view to obtaining the degree of  
Doctor of Economics of the European University Institute

### **Examining Board**

Andrea Mattozzi, EUI (supervisor)

David Levine, EUI

Alfredo Di Tillio, Università Bocconi

Dino Gerardi, Collegio Carlo Alberto, Università degli Studi di Torino

© Lorenzo Verstraeten, 2019

No part of this thesis may be copied, reproduced or transmitted without prior  
permission of the author





### Researcher declaration to accompany the submission of written work

I <Lorenzo Verstraeten> certify that I am the author of the work <Essays in Microeconomics> I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297)).

The copyright of this work rests with its author. [quotation from it is permitted, provided that full acknowledgement is made.] This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

### Statement of inclusion of previous work (if applicable):

I confirm that chapter <2-3> was jointly co-authored with <Ms Julie Pinole> and I contributed <50%> of the work.

### Signature and Date:

Firenze, 29/04/2019

Lorenzo Verstraeten



# Essays in Microeconomics

Lorenzo Verstraeten

February 2019





# Abstract

This dissertation consists of three self-contained essays in microeconomics.

The first chapter studies a principal-agent model where a biased agent can costly collect information useful for the principal. I study what is the optimal contract the principal should commit to, when she cannot do contingent transfers to the agent. When the agent's value of information is higher than its cost, the optimal mechanism is a threshold delegation rule. The principal allows the agent to choose among all the available actions up to some threshold. This threshold is increasing in the parameter measuring the cost of information. Otherwise, the principal will commit to extreme biased behavior to induce information acquisition. The utility of the principal is non-monotonic in the cost of information. While inducing information acquisition becomes more difficult with higher cost, certain deviations in the acquisition stage become more expensive and thus less profitable for the agent.

The second chapter is coauthored with Julie Pinole. Knowing that Individuals interact with their peers, we study how a social planner can intervene, changing these interactions, in order to achieve a particular objective. When the objective is welfare maximization, we describe the interventions for games of strategic complements and strategic substitutes. We show that, for strategic complements, the planner uses resources to target central players; while she divides individuals into separated communities in the case of strategic substitutes. We study which connections she targets in order to achieve these goals.

The third chapter is coauthored with Julie Pinole and analyzes a model of contagion on social network. We ask how a social planner should intervene to prevent contagion. We characterize the optimal intervention and the cost associated. We discuss the intuition behind the choice of the planner and we provide comparative static on the cost of intervention for different type of network.

# Acknowledgments

I am indebted to my advisor Andrea Mattozzi for his guidance. I would like to thank my coauthor Julie Pinole and all the participants to the EUI working group for helpful comments and discussions.

# Contents

<b>Chapter 1: Optimal Contracts with No Transfers and Costly Information Acquisition</b>	<b>1</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Literature Review</b>	<b>5</b>
<b>3 The Model</b>	<b>7</b>
<b>4 The Problem of the Principal</b>	<b>9</b>
4.1 $c \in [0, \frac{1}{12}]$ , Low Cost of Information . . . . .	10
4.2 $c > \frac{1}{12}$ , High Cost of Information . . . . .	12
4.2.1 $b > 0$ . . . . .	16
<b>5 Welfare Analysis</b>	<b>19</b>
5.1 $b = 0$ . . . . .	22
<b>6 Conclusions</b>	<b>23</b>
<b>Chapter 2: Optimal intervention for network games</b>	<b>36</b>
<b>1 Introduction</b>	<b>36</b>
<b>2 Literature Review</b>	<b>38</b>
<b>3 The Model</b>	<b>40</b>
3.1 The setup . . . . .	40
3.2 Equilibrium . . . . .	40
<b>4 Closest network structure to implement a chosen vector of actions</b>	<b>41</b>
4.1 Case 1: Unconstrained intervention . . . . .	42
4.2 Case 2: Symmetric intervention . . . . .	44
4.3 Case 3: Sparse intervention . . . . .	45
<b>5 Closest network structure that maximizes welfare</b>	<b>46</b>
5.1 The planner's problem . . . . .	46

5.2	Equilibrium profile's reaction to planner's intervention . . . . .	46
<b>6</b>	<b>Comparison with GGG</b>	<b>52</b>
<b>7</b>	<b>Network structure analysis</b>	<b>54</b>
<b>8</b>	<b>Other Interventions</b>	<b>58</b>
<b>9</b>	<b>Appendix</b>	<b>59</b>
9.1	Proof of theorem 1 . . . . .	59
9.1.1	Projection of the equilibrium profile before intervention . . . . .	59
9.1.2	Projection of the equilibrium profile after intervention . . . . .	61
9.1.3	Comparing the ratio of projections on two different right singular vectors . . . . .	61
9.1.4	Proof of Corollary (2) . . . . .	63
9.2	Proof of proposition 9 . . . . .	63
9.3	Optimal intervention under symmetry constraint . . . . .	67
	<b>Chapter 3: Stopping contagion: optimal network intervention</b>	<b>69</b>
<b>1</b>	<b>Introduction</b>	<b>69</b>
<b>2</b>	<b>Literature Review</b>	<b>72</b>
<b>3</b>	<b>The Model</b>	<b>74</b>
3.1	Discussion of the epidemic threshold . . . . .	75
<b>4</b>	<b>Optimal immunization</b>	<b>78</b>
<b>5</b>	<b>Applications</b>	<b>82</b>
5.1	Example 1 . . . . .	83
5.1.1	Initial structure . . . . .	83
5.1.2	Post-intervention structure . . . . .	84
5.1.3	Comparative statics of the cost of intervention . . . . .	84
5.2	Example 2 . . . . .	87
5.2.1	Initial structure . . . . .	87
5.2.2	Post-intervention structure . . . . .	88

5.2.3	Comparative statics of the cost of intervention . . . . .	88
5.3	Example 3: homogeneous versus heterogeneous networks . . . . .	90
<b>6</b>	<b>Conclusion</b>	<b>92</b>



# Chapter 1

## Optimal Contracts with No Transfers and Costly Information Acquisition

### 1 Introduction

Information is crucial in decision making. Typically, information is dispersed among various parties and it is costly to acquire. In the first case, an informed agent, with different preference with respect to a principal, might try to mislead her by misreporting information. I call this a problem of “adverse selection in report”. In the second case, an agent, delegated by a principal might exert too little effort in information acquisition. I call this a problem of “moral hazard in information acquisition”.

One of the most interesting and studied situation where these two problems arise is the relation between the Government and the Bureaucracy. The problem of “adverse selection in report” might emerge because the ideology of the Government in charge and the one of an agency are misaligned. Clinton & al. (2011), in an empirical study, documents US governmental agencies bias, identifying which agency can be considered liberal and which one conservative. Several reasons might explain the difference in ideology of various agencies. Some agencies might be created by a Democratic President, other by a Republican one. The appointees in each agency might have different political vision. Agencies with different mission, such as the EPA or the Defense might attract more liberal or more conservative employees. Instead, the problem of “moral hazard in information acquisition” might be caused by the price of information. Even if it is difficult to quantify “research effort” scholars (See Stephenon (2011) for example) point out that Boreaucrats might invest time and resources to improve decision making. Data collection, study of academic literature, consultation of outside parties, implementation of pilote projects are all examples of how an agent can access to information by paying a price.

The principal can offer a performance-based contract in order to mitigate these two kind of problems. Rewarding the agent for correct and precise information is a way for the principal to soften both the adverse selection in report and the moral hazard in information acquisition. There are, however, several environments where this is not possible. Policies might have long-term outcome making it difficult to reward an agent. Finally, it might be difficult to measure the benefit of a policy for each individual in the society and charging

him for the cost accordingly. In fact, what we observe (specially in the public sector) is that wages are mostly fixed and independent of policy outcome.

In this paper I analyze the optimal mechanism to which a principal should commit to, when contingent transfers are ruled out. I describe which channels a principal might use to incentivize correct and precise information acquisition and transmission in the context of a relatively standard principal-agent framework.

When contingent transfers are ruled out the only way to influence the decision of the agent for the principal is to use the incentives of the agent in the project he is evaluating. This implies that, there is not advantage in keeping the decision power for the principal. In order to incentivise the agent for correct decision making, the principal should simply select the appropriate action space from which the agent will choose. The problem I analyze reduces to the choice of a menu of available actions that the principal offers to the agent.<sup>1</sup>

At the beginning of the game the principal decides which actions the agent can and cannot take. The agent, knowing this, decide how much and which type of information to buy. After that, he selects the action that maximizes his expected payoff from the action space. When the principal selects the actions available to the agent, she does it, taking into consideration both the adverse selection problem due to the conflict of interest and the moral hazard problem due to the cost of information acquisition. It is worth noting that the presence of a conflict of interest between the two players, combined with the cost of information, not only influence the action choice of the agent (once the uncertainty is realized) but it critically shapes the type of information the agent acquires.

Specifically, the agent can partition the state space into intervals and learn in which of them the State of the World falls. I call each partition an investigation. The cost of each investigation is the product of a parameter measuring the expertise of the agent performing the investigation times the decrease of uncertainty associated to the investigation. I assume that the more informative an investigation is, the more the agent will have to pay. Furthermore, agents differ in their expertize to perform the investigation. An agent is more expert with respect to an other one if he is able to perform any investigation at a lower cost. An alternative interpretation is to think of the cost of an investigation obtained as a parameter measuring how difficult is the problem we want to solve times the decrease of uncertainty associated to the investigation.

If the principal could buy herself information, at the same cost as the agent, she would buy an investigation as long as the expected value of information would be smaller than the cost. I say that an agent is one that

---

<sup>1</sup>For a similar argument see Holmstrom (1978), Green & Stokey (1981), Melumad & Shibano (1991) and Szalay (2005).



has *low cost* of information acquisition when this condition is satisfied. Alternatively I can think that we are facing a problem that requires information at low cost. I show that, as long as the cost of information is *low* the optimal contract is a threshold delegation rule. The principal let the agent take any action into a certain interval, whose bound are a function of the cost of information. The higher is the cost, the more the principal decides to delegate (make the bounds less binding). The principal rewards the agent, for information acquisition, giving more decision power, letting the agent make better use of the information he learns.

There are several situations where we can observe the kind of (threshold delegation) rule I obtained. In the Government-Bureaucrat relation Gailmard (2008) documents how, in the U.S., the legislature can create a "window of discretion" for agencies to select policies. Examples are the Transportation Act of 1920, delegating the Interstate Commerce Commission to set rail rates but only for specific routes or the Clean Air Act Amendments of 1970 giving the EPA the possibility to set limits on pollutants but with thight restrictions different from region to region. <sup>2</sup>

When the cost of information is *high* I ask whether it is possible to induce information acquisition. If the principal would have to buy herself information she would not do it, as the value would not cover the cost. Given that the cost of information is bared by the agent, the principal might nevertheless want to induce information acquisition. First, I show that the principal can induce only a reduced amount of information acquisition. The idea is that the moral hazard problem in information acquisition is now very severe, making underinvesment in information acquisition difficult to rule out. Finally, I show that the optimal choice for the principal is to make the agent choose among two extreme actions (far from the expectation of the prior).

The agent, facing two very risky choices, will be better off acquiring information. By removing a "safety net" to the agent the principal forces him to invest in acquisition. When there is no conflict of interest between the two parties the principal makes the agent choose among two symmetric actions. In this case there exists only a problem of "moral hazard in acquisition". The more accurate the information that the agent buys the better it is for the principal. When there is a conflict of interest between the players the optimal rule is different. The principal makes the agent chooses among two extreme actions that are biased in favor of the agent's preference. Counter intuitively, the principal might not induce the agent to buy very accurate information. In fact, it is possible that the principal "promotes favorable use of information ex-post

---

<sup>2</sup>High level manager often delegate to lower levels' ones but setting some limits on investment or hiring decisions. Judges are often constrained in the lenght of the sentence they choose.

sacrificing efficient acquisition ex-ante”<sup>3</sup>. An high cost makes information acquisition difficult (the moral hazard problem is severe) on one side, but could make some deviations, during the acquisition phase, less profitable and therefore mitigate the adverse selection problem.

The paper is organized as follows: Section 2 compare the paper with the related literature. Section 3 lays out the basic model. Section 4 presents the main results. Section 5 generalizes some of the results. Section 6 concludes.

---

<sup>3</sup>Stephenson(2011)

## 2 Literature Review

Crawford and Sobel (1984) ask what are the equilibria of a communication game where an informed agent can send messages to an uninformed principal taking an action affecting the payoffs of both players. Melumad & Shibano(1991) and Goltsman & al. (2009) study what happens in the Crawford and Sobel (1984) cheap talk's model when the principal has commitment but cannot do contingent transfers to the agent. In their model, as in Crawford and Sobel(1984) information is exogenously given. The agent enters in the game knowing what is the State of the World. The principal select a menu of actions from which the agent takes his choice. They found that the principal optimally chooses a threshold delegation rule. She allows the agent to take any action in a convex interval with bounds dependent on the parameter measuring the misalignment of preferences between the two players. My model is a generalization of theirs, as information is expensive to acquire. When the cost of the cost of acquisition goes to zero the results I obtain coincide with theirs. The principal, in the game studied by Melumad & Shibano(1991) and Goltsman & al. (2009), using her commitment power, tries to solve the problem of adverse selection caused by the conflict of interest between her and the agent. The additional difficulty present in my model is that the principal will have to tackle also the moral hazard problem caused by the price of information.

Recently, Di Pei (2015), Argenziano & al (2016) and Deimen & Szalay (2017) endogenize information acquisition in communication. They study cheap talk models with costly information acquisition. The three models differ in the technology that the agent can use to collect information. The peculiarity of the technology used to collect information in the model of Di Pei (2015) is that the agent can “address” search effort in the state space. For example, he can choose to have very detailed information when the State of the World is extreme and only a low informative signal when the State of the World is moderate. I use the framework of Di Pei (2015) to study how a principal with commitment and without the possibility of doing contingent transfers should optimally delegate to the agent.

Szalay (2005) studies the problem of how a principal can induce information acquisition by an agent. He found that commitment to extreme behavior by the principal can induce information acquisition. Reducing the number of choices for an agent can induce him to invest in information. The model studied by Szalay (2005) differs from mine in, at least, two dimension. First, in Szalay (2015) there is no conflict of interest between the players. Only the moral hazard component is present. Finally, the technology used to collect information by the agent is different. These two different assumptions imply different results. First, in the optimal contract, in Szalay (2015), the agent choose among a continuous of actions, while I show that for

high cost the agent will face a binary (Low or High) choice. Second, in Szalay (2005), the principal never offers a threshold delegation contract, as it happens in my model when the cost of information is low. Finally, given the absence of the adverse selection problem, in Szalay (2005), a lower cost of information is always preferred by the principal.

The paper contributes to the large literature on mechanism design with endogenous information. Angelucci (2017) studies the interesting problem of how to incentive two different agents to collect information in a model similar to Szalay (2015). Li (2001), Gerardi & Yariv (2008) Cai (2009), Gershkov & Szentes (2009) study mechanism design with endogenous information in collective decision-making.

The economics literature has focused on methods to solve the adverse selection problem alone or the moral hazard problem alone, not considering how these two aspects combine together. I show how using only one instrument to solve the two problems results in “a inevitable trade-off between promoting efficient use of information ex-post and stimulating efficient acquisition ex-ante”.<sup>4</sup>

Referring to this problem Sobel (2011) says “I am unaware of studies that characterize optimal institutions for acquisition and transmission of endogenously generated information for the model ... This is a natural question for future research.”

The following diagram is an (incomplete) picture of the closer works.

		Information cost →	
← Commitment Power	cheap talk	$c = 0$	$c > 0$
		Crawford, Sobel(1982)	Di Pei(2015)
	contract	Melumad, Shibano(1991)	Argenziano & al.(2016), Deimen,Szalay (2017)
		Goltsman & al.(2009)	My model
			Szalay (2005), Angelucci(2017)

---

<sup>4</sup>Stephenson (2011)

### 3 The Model

In this section I present the model.

The state of the world is denoted by  $\theta$  and it is drawn from a uniform (commonly known) distribution  $\theta \sim U[-\frac{1}{2}, \frac{1}{2}]$ . When action  $a \in [-\frac{1}{2}, \frac{1}{2}]$  is chosen in state  $\theta$  the Principal gets utility  $U(a, \theta) = -(a - \theta)^2$  and the Agent gets  $V(a, \theta) = -(a - \theta - b)^2$ , where  $b \in [0, \frac{1}{8}]$ <sup>5</sup> measures the misalignment of preferences between the two players. The Principal wants to match the state while the agent has some bias towards high actions.

The agent has the possibility to buy an interval partition<sup>6</sup>  $\psi \in \Psi$  of the state space  $[-\frac{1}{2}, \frac{1}{2}]$ . When the agent buys  $\psi \in \Psi$  he learns in which subset of  $\psi$  the state falls. For finite partition we have

$$\psi = \left\{ [\theta_0, \theta_1], (\theta_1, \theta_2], \dots, (\theta_{n-1}, \theta_n], \text{ with } -\frac{1}{2} = \theta_0 \leq \theta_1 \dots \leq \theta_n = \frac{1}{2} \right\} \quad (3.1)$$

The agent learns that  $\theta \in (\theta_{i-1}, \theta_i]$  after the acquisition of the information partition. The acquisition stage is not observable by the principal and the total utility of the agent is given by the payoff from the action minus the cost of buying the partition.

**Assumption 1.** *The cost to buy a partition  $\psi$  is*

$$c(\psi) = c \left( 1 - \sum_{\Theta \in \psi} \mu_0(\Theta)^3 \right), \quad (3.2)$$

where  $\mu_0(\Theta)$  is the Lebesgue measure of the interval  $\Theta$  belonging to the partition  $\psi$ .<sup>7</sup>

The previous expression is based on the assumption that a more informative partition is more expensive to acquire. The following remarks and the discussion in Appendix B elaborates on this.

To fix ideas, imagine that an agent choose an  $n$  - *elements* partition of the state space  $\psi$ . Then  $c(\psi) = c \left( 1 - \sum_{i=0, \dots, n} (\theta_i - \theta_{i-1})^3 \right)$ . After the acquisition of the partition  $\psi$  the agent's posterior would be  $\theta \sim U[\theta_i - \theta_{i-1}]$ , with probability  $(\theta_i - \theta_{i-1})$ . Each posterior distribution has as variance  $\frac{1}{12}(\theta_i - \theta_{i-1})^2$ .

<sup>5</sup>I consider the range of bias  $b$  for which in Crawford & Sobel (1982) there exists a non-bubbling equilibrium.

<sup>6</sup>Note that, differently from Di Pei (2015) we allow the agent to buy partitions made of infinite elements. This implies that, in particular, the agent has to possibility to learn the exact state. A partition of the set  $[-\frac{1}{2}, \frac{1}{2}]$  is a set of nonempty subsets such that every element  $\theta$  in  $[-\frac{1}{2}, \frac{1}{2}]$  is in exactly one of these subsets and the union of these subsets gives  $[-\frac{1}{2}, \frac{1}{2}]$ .

<sup>7</sup>The definition of the sum over an uncountable set is an abuse of notation.

Note, however that the sum would be always  $\leq 1$  and it can be  $\mu_0(\Theta) > 0$  only for a finite number of indices  $\Phi$ .  
Formally  $c(\psi) = c \left( 1 - \int_{\psi} \left( \int_{[-\frac{1}{2}, \frac{1}{2}]} 1_{\Phi} d\mu_0 \right)^3 d\phi \right)$ .

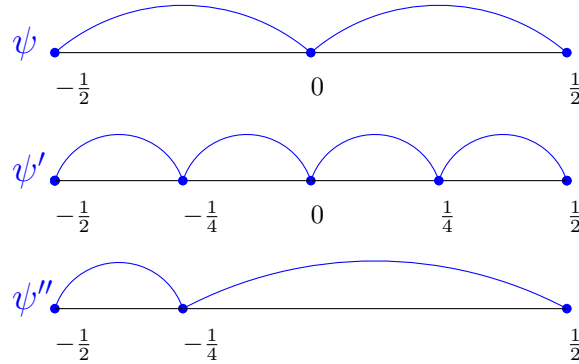
Hence, given  $\psi$ , the expected variance of the posteriors is given by  $\sum_i (\theta_i - \theta_{i-1}) \frac{1}{12} (\theta_i - \theta_{i-1})^2$ . When the agent has no access to information his posterior would be equal to his prior  $\theta \sim U[\theta_i - \theta_{i-1}]$  so that the expected variance of the posterior (= prior) is given by  $\sum_i \left(\frac{1}{2} - \left(\frac{1}{2}\right)\right) \frac{1}{12} \left(\frac{1}{2} - \left(\frac{1}{2}\right)\right)^2 = \frac{1}{12}$ . Lower expected variance in the posteriors is associated to higher cost.<sup>8</sup>

The choice to take a cubic cost function<sup>9</sup> is particularly appealing in this setting as it turns out that if the problem would be a single-agent problem, ie the principal would buy himself information, the cost of an information partition would be proportional to its value. We can, in fact, in this case compute the value of a partition  $\psi$  as the expected utility the principal gets from it minus the expected value she would get without information, that is:

$$\max_{a_1, \dots, a_n} \left[ -\sum_{i=1}^n \int_{\theta_{i-1}}^{\theta_i} (a_i - \theta)^2 d\theta \right] - \max_x \left[ -\sum_{i=1}^n \int_{-\frac{1}{2}}^{\frac{1}{2}} (x - \theta)^2 d\theta \right] = -\frac{1}{12} \left( \sum_{i=1}^n (\theta_i - \theta_{i-1})^3 - 1 \right). \quad (3.3)$$

In such a setting it should be clear that the principal would invest in information acquisition (acquiring the finest possible partition of the state space) as long as  $c \leq \frac{1}{12}$  while no information should be acquired when  $c > \frac{1}{12}$ . The point  $c = \frac{1}{12}$  is where an hypothetical single-agent would be indifferent between buying information or not. In the analysis that follows I will say that the cost of information is low if  $c < \frac{1}{12}$  and that the cost of information is high if  $c > \frac{1}{12}$ .

In order to get an idea of how we compare two different information partitions consider the following figure:



<sup>8</sup>Other choices of the cost function would be coherent with the notion of informativeness in terms of Blackwell theory. Check Appendix B for a further discussion.

<sup>9</sup>Check Di Pei (2015) for a further discussion.

FIG.1 The cost function

It should be clear that  $c(\psi) < c(\psi')$  and  $c(\psi') > c(\psi'')$ . The assumption on the cubic cost function rank also  $c(\psi) < c(\psi'')$  as the expected variance of the posteriors' distributions is lower in  $\psi''$  wrt  $\psi$ .

The way we should interpret the ranking between  $\psi$  and  $\psi''$  is that it is more informative a signal  $\psi$  that, with the same probability, gives an accurate measure of the state wrt one ( $\psi''$ ) that gives with a low probability a very precise information and with an high probability a non-precise information. In Appendix B I show that this is consistent with a large class of cost function and I show what happens to some of the main results when I remove this assumption

Finally, I assume that it is not possible for the principal to do transfers to the agent.

Before the agent invests in information acquisition the principal has the possibility to commit to a mechanism. That is, she specifies a message space  $M$  such that, if the agent choose  $m \in M$  she plays  $a(m)$ .

### Timing

1. Principal chooses  $M$ , and for any  $m \in M$  commit to  $a(m)$ ;
2. Agent buys  $\psi \in \Psi$  and learns that  $\theta \in \Theta$ ;<sup>10</sup>
3. Agent chooses  $m \in M$  and the Principal perform  $a(m)$ ;
4. Payoffs are realized.

## 4 The Problem of the Principal

Given the nature of the problem and the impossibility of doing transfers; the principal cannot do better than choosing a subset of the action space from which the agent will select an action.<sup>11</sup> There is, in fact, no gain in keeping decision-power for the principal.<sup>12</sup> Hence we can write her problem as follows:

Let  $\Gamma \in \Gamma$  be any possible closed subset of  $[-\frac{1}{2}, \frac{1}{2}]$ .

. The Principal solves:

---

<sup>10</sup>Only pure-strategy in the acquisition stage

<sup>11</sup>Check Holmstrom (1978,1984), Green and Stokey (1981), Melumad & Shibano (1991) and Szalay (2005) for a similar argument and Appendix A for a formal argument.

<sup>12</sup>In the timeline, in the previous section, after 1. the principal is a passive player.

$$(P) \quad \max_{\Gamma \in \Gamma} - \sum_{\Theta \in \psi} \int_{\Theta} (a - \theta)^2 d\theta \quad s.to. \quad (4.1)$$

$$i) \quad a(\Theta) \in \arg \max_{a \in \Gamma} - \int_{\Theta} (a - \theta - b)^2 d\theta$$

$$ii) \quad \psi \in \arg \max_{\psi' \in \Psi} - \sum_{\Theta \in \psi'} \int_{\Theta} (a(\Theta) - \theta - b)^2 d\theta - c(\psi')$$

The principal maximizes her expected utility by choosing which actions to make available to the agent considering two conditions. The first is that the agent, knowing that the state belongs to  $\Theta$ , will choose the action that maximizes his expected utility among the actions available in  $\Gamma$ . This mimics a classical Incentive Compatibility constrain and it refers to what I call “adverse selection problem in report”. The second condition is that the agent buys the interval partition that maximizes his expected utility minus the cost of the partition. This constrain is the one related to the moral hazard component of the problem.

#### 4.1 $c \in [0, \frac{1}{12}]$ , Low Cost of Information

First, I want to analyze what happens when the cost to collect information is low.<sup>13</sup>

Two benchmark cases can help me designing the optimal contract:

- The first is the case  $b = 0$ . The principal, not having any conflict of interest with the agent, should leave complete freedom to him. The agent, being the cost of information low will buy a “complete” partition of the state space, he will learn  $\theta$  and plays  $a = \theta$ . The principal would get the most informative information partition chosen by the agent and the action that goes with it, while the agent is willing to buy it as the cost of it is lower than the benefit he gets by choosing an action based on a more precise posterior.
- When  $c = 0$ , Melumad & Shibano (1991) provides a solution for the problem for any  $b > 0$ . In fact, they show that, when the agent knows what is the state, the principal should allow the agent to take any action below  $\frac{1}{2} - b$ . The best way to use the information owned by the agent is, in fact, to delegate to him for low state states and keeping decision power for high states. The agent will choose an action  $a = \theta + b$  as long as  $\theta \leq \frac{1}{2} - 2b$  while he will have to choose  $a = \frac{1}{2} - b$  when  $\theta > \frac{1}{2} - 2b$ . The following figure illustrates how we can think of this problem in terms of information acquisition: The principal chooses the action space  $\Gamma = [-\frac{1}{2}, \frac{1}{2} - b]$ . The agent learns the exact state for any  $\theta \leq \frac{1}{2} - 2b$ , playing  $a = \theta + b$ ; or he learns that the SoW is between  $\frac{1}{2} - 2b$  and  $\frac{1}{2}$  and plays  $a = \frac{1}{2} - b$ .

---

<sup>13</sup>The cost of information is considered to be low if the cost is lower with respect to price of information in terms of the Remark II.2



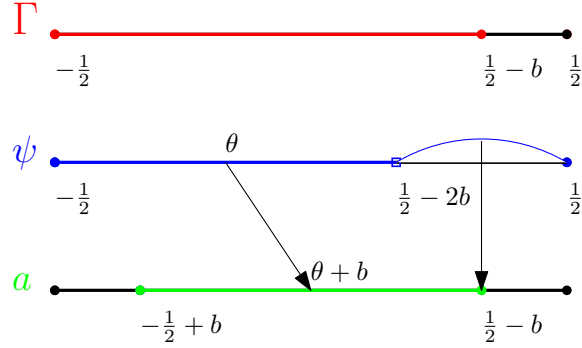


FIG.2 Optimal rule  $c = 0$ ,  $b > 0$

I will build on this two specific cases a complete characterization of the optimal contract. The following proposition characterize the optimal mechanism for any  $b \geq 0$ ,  $c \in (0, \frac{1}{12}]$  :

**Proposition 1.** *If  $c \in (0, \frac{1}{12}]$  there exists a unique  $\alpha(c)$*

$$\Gamma = \left[ -\frac{1}{2}, \alpha(c) \right]$$

where  $\alpha'(c) > 0$  for any  $c \in [0, \hat{c}]$  and  $\alpha(c) = 1$  for  $c \in [\hat{c}, \frac{1}{12}]$ .

Proof in the Appendix

In order to get some intuition behind this result let me start by observing what happens for low state. If the principal leaves the decision to the agent he will learn the exact state  $\theta$  and he will play  $\theta + b$ . This is because as long as  $c \leq \frac{1}{12}$  and he can play “his favorite action” he has an incentive to buy as much information as possible. This is exactly what was happening in the case of  $c = 0$  described by Melumad & Shibano (1991). The main difference will be for high states of the world. In fact, if, as in the case  $c = 0$ , I leave the agent chooses all the actions up to  $\frac{1}{2} - b$  he will NOT buy exact information below  $\frac{1}{2} - 2b$ . He will instead buy a less informative partition  $\left\{ \cup_{x \leq \frac{1}{2} - 2b - \varepsilon} \{x\}, [\frac{1}{2} - 2b - \varepsilon, \frac{1}{2}] \right\}$ , for some  $\varepsilon > 0$ , to play  $\frac{1}{2} - b$ .

The fact that information is costly creates this additional “moral hazard” problem for the principal. She will have to trade off this, with the adverse selection problem. Hence, the principal has to decide whether to delegate more (or less) to the agent inducing more (or less) information acquisition. The problem boils down to the search of the last action that she allows the agent to take. The trade-off is solved in favor of more delegation (solving the “moral hazard” and keeping the “adverse selection”).

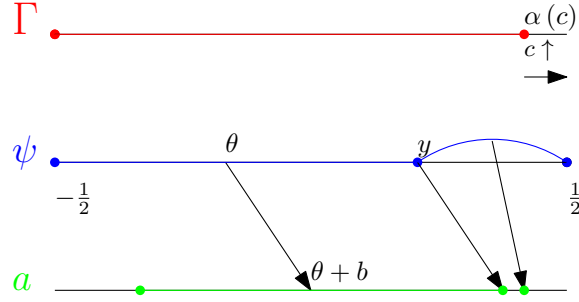


FIG. 3 Optimal rule  $c > 0$ ,  $b > 0$

The principal prefers to leave the agent decision power (by letting him choose higher actions when the cost increases) in order to compensate for the expenditure in information acquisition. When the cost of information acquisition is large the principal will allow the agent to take any action in the action space. The agent will buy exact information up to some value  $y$  and learns  $\theta$  to play  $\theta + b$  or learns that the SoW is in  $[y, \frac{1}{2}]$  and play  $\alpha(c)$ .

The optimal rule found in Proposition 1 takes the name of “threshold delegation rule” in the literature. The difference with most of the settings<sup>14</sup> that have been studied is that the information is endogenously acquired in this model and therefore the rule depends on the cost of the information acquisition. In the empirical I.O. literature we find several examples of top managers delegating using this kind of rule.

An important observation is that, for this range of the costs, the main problem for the principal is the “adverse selection” one. In fact, in case  $b = 0$ , he would be able to get the first best<sup>15</sup> no matter what is  $c$ , while for  $b > 0$  she loses something (wrt the case  $c = 0$ ) only for high states.

## 4.2 $c > \frac{1}{12}$ , High Cost of Information

I will first provide a lemma to characterize the number of actions that the principal can induce the sender to take.

**Lemma 1.**     • If  $c > \frac{1}{12}$  the principal can induce up to 2 different actions (or equivalent a partition with a maximum of 2 elements);

• If  $c > \frac{1}{3}$ ,  $\Gamma = \{0\}$ .

<sup>14</sup>Check Alonso & Matouschek (2007) for a detailed discussion.

<sup>15</sup>in the sense that ruling out contingent transfers would do not reduce do not reduce the utility of the principal.

Proof in the Appendix

Note that we could expect that, when  $c > \frac{1}{12}$ , there would not be space for the principal to induce the agent to buy information. In fact, a principal that could acquire information would not do it herself when  $c > \frac{1}{12}$ . This should already suggest that the possibility to induce information acquisition is limited. The lemma specifies how binding is this limit.

The idea for the second bullet is that when the cost of collecting information is very high then the agent will never buy any. The only option for the principal in this case is to choose the action that goes with the expectation of the prior, i.e.  $a = 0$ .

In order to get some intuitions for the first statement let me assume, by contradiction, that the principal could induce 3 actions and let me show that it exists a profitable deviation for the agent.<sup>16</sup>

WLoG<sup>17</sup> suppose that the principal wants to induce two negative actions and a positive one. We should be able to find  $-\frac{1}{2} \leq a_1 < \theta_1 < a_2 < \theta_2 < a_3 \leq \frac{1}{2}$ , with  $a_2 < 0$ , for some  $a_1, a_2$  and  $a_3$  chosen by the principal and for some choices  $\theta_1$  and  $\theta_2$  of the agent.

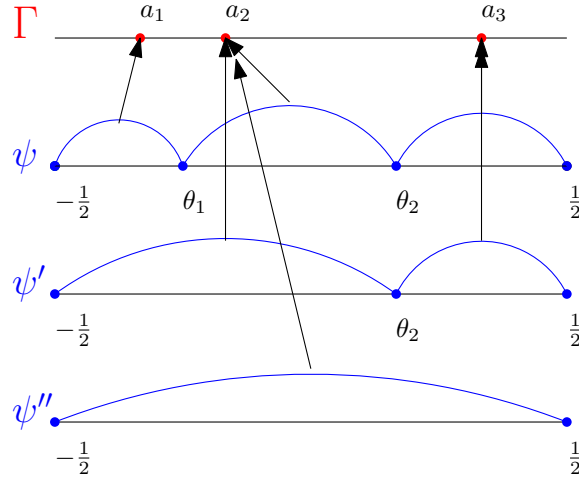


FIG. 4 The principal has to rule out deviation  $\psi'$  and  $\psi''$

As in the figure, the principal would pick an action set  $\Gamma$  with 3 elements and the agent would partition the state space according to some  $\psi$ .

<sup>16</sup>Meaning that the agent would buy an information partition of 3 elements and he would take one different action for any of the 3 different posteriors.

<sup>17</sup>it is easier to show the existence of a profitable deviation when she wants to induce 3 negative (or positive) actions.

The first type of deviation ( $\psi'$ ) we consider is one where the agent buys  $\{[-\frac{1}{2}, \theta_2]; [\theta_2, \frac{1}{2}]\}$  and plays  $\arg \min \left\{ \left\| \frac{-\frac{1}{2} + \theta_2}{2} - a_1 - b \right\|; \left\| \frac{-\frac{1}{2} + \theta_2}{2} - a_2 - b \right\| \right\}$  or  $a_3$ .

In order to rule out this deviation we need to impose a condition on the distance between  $a_1$  and  $a_2$ . When the two actions available for the agent are close one to the other, he will not be able to use efficiently the information he bought. In fact, after learning (at very high cost) in which element of the partition the state is, it will not make a large difference to choose  $a_1$  instead of  $a_2$  if they are close one to the other. Hence, he prefers to pull the first the two elements of the partition together. Now, given that  $a_2$  has to be distant enough from  $a_1$ , an other deviation ( $\psi''$ ) becomes profitable.

The agent prefers not to buy information ( $\psi''$ ) and plays  $a_2$ , being this action very closed to the one that maximizes his utility under the prior distribution.

What the previous lemma implies is that we will only have to look for 2 actions in order to maximize the principal expected utility. This makes the maximization problem of the principal much easier to deal with.

Before solving for the optimal mechanism, I state the following proposition that characterizes the rule to which the principal should commit when there is no conflict of interest between the two parties.

This intermediary result would help me to get some insights on the general rule ( $b > 0$ ) and to do a comparison with the results available in the literature.

**Proposition 2.** *If  $b = 0$ , the principal chooses:*

$$\Gamma = \begin{cases} \{-\frac{1}{4}, \frac{1}{4}\} & \text{if } c \in (\frac{1}{12}, \frac{1}{6}] \\ \{-y, y\} & \text{if } c \in (\frac{1}{6}, \frac{1}{3}] \end{cases}, \text{ where } y = \frac{3}{2}c.$$

I can guess and verify the solution. If the principal could induce the two actions  $-\frac{1}{4}$  and  $\frac{1}{4}$  under the condition that the signal bought by the agent is  $\{[-\frac{1}{2}, 0]; [0, \frac{1}{2}]\}$  we would have the optimal mechanism. Solving the maximization we can verify that this is the case as long as  $c < \frac{1}{6}$ .

In  $c = \frac{1}{6}$  the agent would be indifferent between buying  $\{[-\frac{1}{2}, 0]; [0, \frac{1}{2}]\}$  and playing  $a_1, a_2$  (according to the realization of the signal) or not buying any signal and playing  $a_1$  (or  $a_2$ ).

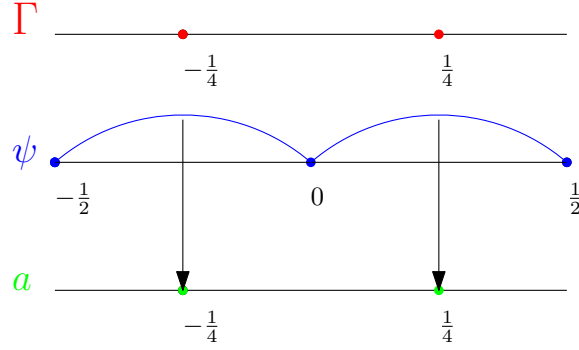


FIG. 5 The optimal rule for  $b = 0$ ,  $c \in (\frac{1}{12}, \frac{1}{6}]$

When  $c$  is larger than  $\frac{1}{6}$  if the principal wants the agent to buy some signal then she has to increase the distance between the two actions  $a_1$  and  $a_2$ . The principal wants to keep the distance between the two actions as small as possible with the condition that the agent keeps buying information. She will keep the agent indifferent between buying and not buying information, making the distance between  $a_1$  and  $a_2$  larger as  $c$  increases.

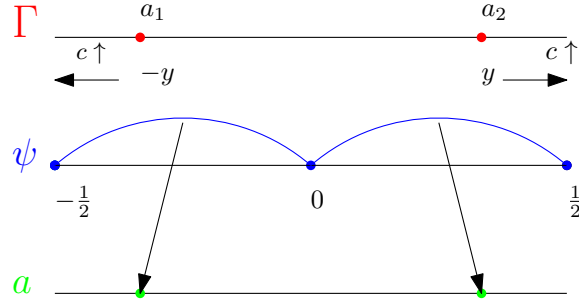


FIG. 6 The optimal rule for  $b = 0$ ,  $c \in (\frac{1}{6}, \frac{1}{3}]$

The idea behind this proposition is that by eliminating intermediary actions (committing to extreme behavior) the principal force the agent to acquire information. This happens because the principal makes the consequences of what the agent chooses more extreme. This result reminds the ones found in Szalay (2005) and Angelucci (2017).

It is worth noting that as long as  $c \leq \frac{1}{6}$  the actions available for the agents allows him to fully use the information he learns. When  $c > \frac{1}{6}$  the agent cannot efficiently use the information he has damaging also the principal. For this range of costs the moral hazard problem is extremely severe and as a consequence we would have mechanism that are ex-post inefficient.

#### 4.2.1 $b > 0$

I want to understand what happens in a situation where the moral hazard problem (due to high cost of information acquisition) interacts with an adverse selection problem (conflict of interest between the agent and the principal). The following Proposition characterizes the optimal rule the principal commits to.

**Proposition 3.** *If  $c \in [\frac{1}{12}, \frac{1}{6}]$ , for any  $b > 0$ , there exists  $c^* \in (\frac{1}{12}, \frac{1}{6}]$  such that the principal chooses  $\Gamma = \{a_1, a_2\}$  with*

$$\begin{cases} a_1 + a_2 > 2b, a_2 - a_1 \leq \frac{1}{2} & \text{if } c < c^* \\ a_1 + a_2 = 2b, a_2 - a_1 = 3c & \text{if } c \geq c^* \end{cases}$$

Proof in the Appendix

In order to get some intuition for this result, I build on the results of Proposition 2.

In the following figure, I show which information partition

$$\psi_1^A = \left\{ \left[ -\frac{1}{2}, \theta_1^A \right], \left( \theta_1^A, \frac{1}{2} \right] \right\} = \left\{ \left[ -\frac{1}{2}, -\frac{b}{1-6c} \right], \left( -\frac{b}{1-6c}, \frac{1}{2} \right] \right\}$$

the agent chooses if the principal offers the same contract  $\Gamma_1 = \{-\frac{1}{4}, \frac{1}{4}\}$  she was offering when  $b = 0$ .<sup>18</sup>

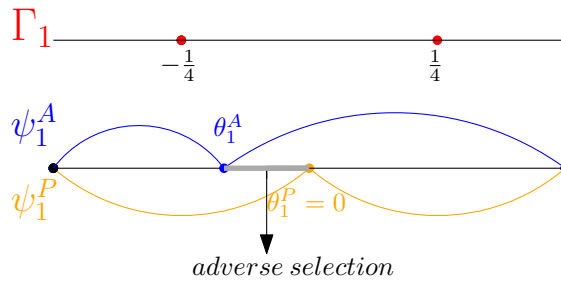


FIG. 7 comparison with contract  $b = 0$ ,

$$\psi_1^P = \left\{ \left[ -\frac{1}{2}, \theta_1^P \right], \left( \theta_1^P, \frac{1}{2} \right] \right\} = \left\{ \left[ -\frac{1}{2}, 0 \right], \left( 0, \frac{1}{2} \right] \right\}^{19}$$

<sup>18</sup>I am considering the cost  $c$  close enough to  $\frac{1}{12}$ .

<sup>19</sup>The principal wants  $\theta^P$  to be the point that is at same the distance between  $a_1$  and  $a_2$

would instead be the information partition the principal would like the agent to buy when the actions available are the ones in  $\Gamma^1$ . The interval  $(\theta^A, \theta^P)$  measures “the adverse selection” component of the problem.

If the principal would “correct” for the bias and offers instead  $\Gamma_2 = \{-\frac{1}{4} + b, \frac{1}{4} + b\}$  the agent would buy the same information partition he was buying in the case where  $b = 0$ . This is the most informative signal that the principal can induce the agent to buy but the principal would still incur in an adverse selection problem measured by the segment  $(\theta_2^A, \theta_2^P)$ .

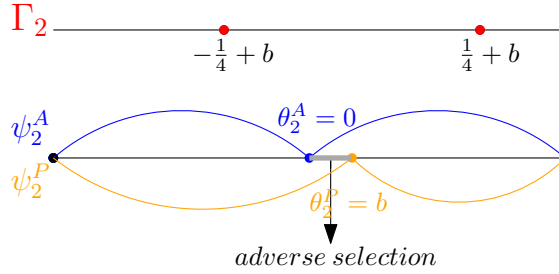


FIG. 8 adjustment for the bias

Proposition 3 tells me that there is a way to reduce this adverse selection problem and do better for the principal. The principal should choose two actions (FIG. 9) that are (on average  $a_1 + a_2 > 2b$ ) even more biased than in  $\Gamma_2$ , in favour of the bias of the agent. This two actions should have lower distance ( $a_2 - a_1 \leq \frac{1}{2}$ ) one from the other wrt to the case where  $b = 0$ .

What the proposition suggests is that the principal should choose two actions that are (on average  $a_1 + a_2 > 2b$ ) more biased in favor of the bias of the agent and with a low distance ( $a_2 - a_1 \leq \frac{1}{2}$ ). If information would have no cost the agent would buy

$$\psi^A = \left\{ \left[ -\frac{1}{2}, \theta^A \right], \left( \theta^A, \frac{1}{2} \right] \right\} = \left\{ \left[ -\frac{1}{2}, \frac{a_1 + a_2}{2} - b \right], \left( \frac{a_1 + a_2}{2} - b, \frac{1}{2} \right] \right\}$$

while the principal would like

$$\psi^P = \left\{ \left[ -\frac{1}{2}, \theta^P \right], \left( \theta^P, \frac{1}{2} \right] \right\} = \left\{ \left[ -\frac{1}{2}, \frac{a_1 + a_2}{2} \right], \left( \frac{a_1 + a_2}{2}, \frac{1}{2} \right] \right\}$$

But, given that the cost of information is increasing in the informativeness of the signal, it is  $c(\psi_1^P) < c(\psi^A)$ . When the principal makes the 2 actions available close ( $a_2 - a_1 \leq \frac{1}{2}$ ) and over-biased ( $a_1 + a_2 > 2b$ ) the value of information becomes smaller for the agent. The agent will want to save on the cost and he will

choose an (cheaper) information partition  $\psi$  closer to the one that the principal prefers. The reduction of the adverse selection problem that the principal can obtain in this way comes at the cost of less information acquisition. In cheap talk models, it might be the case that, a receiver prefers to face a less informed sender. In fact, more information might create larger adverse selection problems. It is similar in this situation where the principal induce less information acquisition to mitigate the conflict of interest.

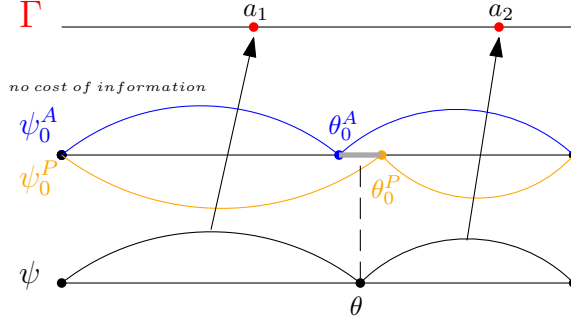


FIG. 9 optimal contract

When the cost of information  $c$  becomes very large the moral hazard problem becomes very severe and, as it was happening in the case  $b = 0$ , the principal will have to set the two actions  $a_1$  and  $a_2$  very distant one from the other. The following corollary specifies the optimal contract for very large values of  $c$ :

**Corollary 1.** *If  $c \in (\frac{1}{6}, \frac{1}{3}]$ , there exists  $c^{**} \in (\frac{1}{6}, \frac{1}{3})$  such that:*

$$\Gamma = \begin{cases} \{-\frac{3}{2}c + b, \frac{3}{2}c + b\} & \text{for any } c \leq c^{**} \\ \{0\} & c > c^{**} \end{cases}$$

The agent will be indifferent between buying or not buying information ( $a_2 - a_1 = 3c$ ) and the contract will look like the one offered in the case where  $b = 0$ . The difference will be that  $a_1$  and  $a_2$  will be biased in favor of the agent and the information partition acquired by the agent would be  $\{[-\frac{1}{2}, b), [b, \frac{1}{2}]\}$ .

When  $c$  becomes even larger, at some point, the principal prefers not to delegate any information acquisition anymore and choose the action that goes with her prior distribution, i.e.  $\Gamma = \{0\}$ .<sup>20</sup>

Differently from the case of low cost of information, the impossibility of doing contingent transfers to the agent affects the utility of the principal both for the adverse selection problem and for the moral hazard one.

<sup>20</sup>It is possible to determine  $c^{**}$  has the cost for which the principal is indifferent between delegating to the agent or taking the action corresponding to the prior expectation.  $\rightarrow -\int_{-\frac{1}{2}}^b (-\frac{3}{2}c + b - \theta)^2 d\theta - \int_b^{\frac{1}{2}} (\frac{3}{2}c + b - \theta)^2 d\theta = -\int_{-\frac{1}{2}}^{\frac{1}{2}} (0 - \theta)^2 d\theta$ .



## 5 Welfare Analysis

In the previous section I characterized the optimal mechanism for all possible cost and bias. In order to make some welfare consideration I first introduce a more general version of the problem. So far, the search for the optimal rule was taking into consideration only the utility of the principal. The agent is, in fact, "forced" to participate to the mechanism without receiving any benefit. If I allow the principal to make flat transfers (not depending on the choice of the agent/realization of the state), I can now impose a participation constraint for the agent that the principal will have to satisfy. If I think that the agent is an employee of an agency and the government (the principal) delegates him a job I can assume that the employee could refuse to take it if he is not compensated to do it. He will incur in the quadratic loss only if he takes the job. His interest in the outcome of the project could be due to career concerns. If he refuses the job he will not have to pay the consequences of a bad policy decision of the government.

The Principal solves:

$$(P_\lambda) \quad \max_{\Gamma', t \in \mathbb{R}^+} -\lambda \sum_{\Theta \in \psi} \int_{\Theta} (a - \theta)^2 d\theta - t \text{ s.to}$$

- i)  $a(\Theta) \in \arg \max_{a \in \Gamma'} -(1 - \lambda) \int_{\Theta} (a - \theta - b)^2 d\theta$
- ii)  $\psi \in \arg \max_{\psi' \in \Psi} -(1 - \lambda) \left[ \sum_{\Theta \in \psi'} \int_{\Theta} (a(\Theta) - \theta - b)^2 d\theta + c(\psi') \right]$
- iii)  $-(1 - \lambda) \left[ \sum_{\Theta \in \psi} \int_{\Theta} (a(\Theta) - \theta - b)^2 d\theta + c(\psi) + t \right] \geq 0$

$\lambda$  measures how much the principal weight the utility from the project wrt the agent  $(1 - \lambda)$ .  $t$  is the wage that the principal can pay to the agent and iii) is the participation constraint of the agent.

The new problem seems to be more realistic. In terms of the examples presented in the introduction,  $t$  can be thought as the budget that the state finances for an agency;  $\lambda$ ,  $1 - \lambda$  represents how much the state and the agency care for the outcome of the investigation.

### IV.1 large $\lambda$

Note that the solution of the problem for  $\lambda$  close to 1 is exactly what we described in the previous section, with the little caveat that the principal will have to pay a wage to the agent that exactly compensate for his expenditure in information acquisition. Being the value for the project infinitely higher for the principal with respect to the agent I should not be surprised that she wants to extract as much as possible from the agent

and repay him through a flat transfer. The following proposition characterize the utility of the principal as a function of the cost of information collection:

**Proposition 4.** *Let  $U$  measure the expected utility of the principal as in  $P_\lambda$ . There exists a  $\lambda^*$  such that for any  $\lambda^* < \lambda < 1$ :*

$$\begin{aligned} \bullet \text{ If } b = 0, \text{ then } \frac{dU}{dc} & \begin{cases} = 0 & \text{if } c \in (0, \frac{1}{12}) \\ = 0 & \text{if } c \in (\frac{1}{12}, \frac{1}{6}) \\ < 0 & \text{if } c \in (\frac{1}{6}, \frac{1}{3}) \end{cases} \\ \bullet \text{ If } b > 0, \text{ then }^{21} \frac{dU}{dc} & \begin{cases} < 0 & \text{if } c \in (0, \frac{1}{12}) \\ > 0 & \text{if } c \in (\frac{1}{12}, c'], \text{ where } c' = \frac{1}{6} + \frac{2}{3}b^2 \\ < 0 & \text{if } c \in (c', \frac{1}{3}] \end{cases} \end{aligned}$$

Proof in the Appendix

The first result, for  $b = 0$ , is quite intuitive. When there is no conflict of interest, the only problem in communication is a moral hazard problem:

1. When information is cheap, i.e.  $c \in (0, \frac{1}{12})$ , both the agent and the principal benefit from information.
2. In  $c \in (\frac{1}{12}, \frac{1}{6})$  the principal induces the acquisition of a two elements partition not depending on  $c$ .
3. When the cost of information is very high, i.e.  $c \in (\frac{1}{6}, \frac{1}{3})$ , the moral hazard problem is very high and the principal has to commit to a mechanism that is ex-post inefficient and this cause a loss of utility when  $c$  increases.

When there exists some bias, i.e.  $b > 0$ , between the two players we have a result that might seem surprising: the utility of the principal is NOT monotonically decreasing in the cost of information acquisition.

---

<sup>21</sup>Note that this derivative does not exist in the points  $c = \hat{c}$ ,  $c = \frac{1}{12}$ ,  $c = c^*$  and  $c = c^{**}$ , defined in Proposition 1 and Proposition 3. Writing  $\frac{dU}{dc}$  is an abuse of notation.

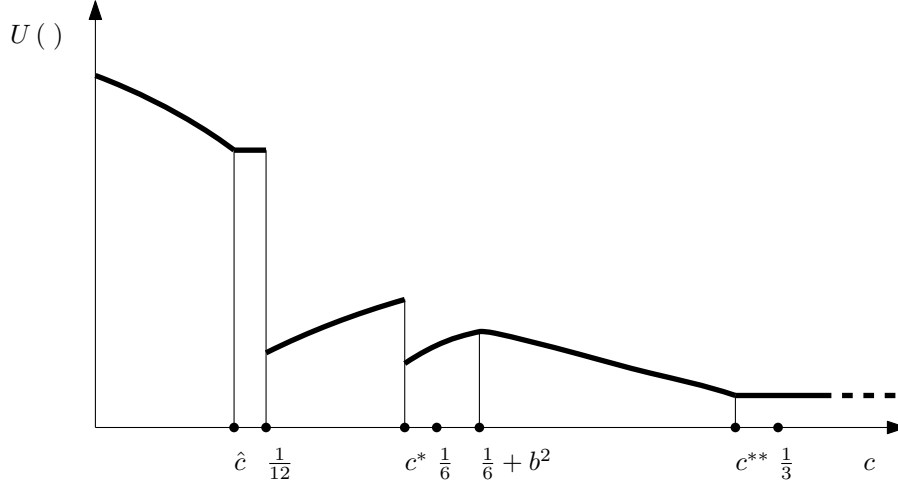


FIG. 10 Non-monotonic effect of the cost on the utility of the Principal

1. When information is cheap, more delegation to the agent (as suggested by the optimal mechanism), as it happens when  $c$  increases, makes the principal worse off.
2. When the price of information is high, the moral hazard problem becomes more severe.<sup>22</sup> At the same time the adverse selection problem becomes less strong as deviations in the acquisition stage become less profitable for the agent. This second effect wins, for a non-empty set of the cost parameters.
3. When the moral hazard problem becomes even more severe the principal is forced to ex-post inefficient contract as in the case of no bias.

This results (in case 2.) suggests that a principal might prefer, in presence of a conflict of interest, to consult an expert that has higher cost of information acquisition.

In order to have a better understanding, consider some  $c > \frac{1}{12}$  small enough and  $b > 0$ . We know that the optimal contract is given by some  $a_1, a_2$  (as described in Proposition 3) and that the agent buys an information partition  $\psi = \{[-\frac{1}{2}, \theta^c], (\theta^c, \frac{1}{2}]\}$ . When  $c$  increases by some small  $\varepsilon > 0$  the principal can still offer the same contract  $\Gamma = \{a_1, a_2\}$  and the agent cannot choose  $\theta^{c+\varepsilon} < \theta$  because it would be too expensive. It must be that  $\theta^{c+\varepsilon} \in (\theta^c, \theta^P)$ . Thus, the principal will be better off. Note that when  $c$  increases the principal might decide (and he does) to change  $a_1$  and  $a_2$ , but in any case she will be better off when the cost of information is  $c + \varepsilon$  wrt  $c$ .

<sup>22</sup>In Di Pei (2015) this aspects also emerge in some dimensions.

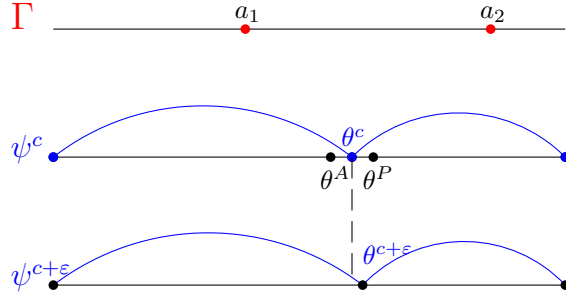


FIG. 11 Increasing utility

### 5.1 $b = 0$

We are not able to characterize the optimal mechanism for any  $\lambda \in (0, 1)$  and  $b > 0$  but we find useful to explore the case where  $b = 0$ . This problem has been studied by Szalay (2005) in an environment where a different types of signals (investigations) could be bought by the agent.

Let denote by  $\Gamma_\lambda$  the solution of problem  $(P_\lambda)$  and by  $\Gamma$  the solution of problem  $(P)$  found in the previous section. The following proposition characterizes the solution  $(P_\lambda)$ , for any  $\lambda \in (0, 1)$ :

**Proposition 5.** *When  $b = 0$ , for any  $c \in [0, \frac{1}{3})$  there exists  $\lambda_c \in (0, 1)$  such that*

$$\Gamma_\lambda = \begin{cases} \Gamma & \text{if } \lambda \geq \lambda_c \\ \{0\} & \text{if } \lambda < \lambda_c \end{cases}$$

and  $\lambda_c$  is increasing in  $c$ .

Proof in the Appendix

If the value of the project for the principal is not so high, she will not ask the agent to collect information when the cost of doing so is large. On the opposite she will enforce information collection when the project is particularly important and the cost is relatively small.<sup>23</sup>

Note that this result does not extend naturally when there is some bias  $b > 0$ . We can, in fact conjecture that the principal might want to change the set  $\Gamma$  according to the value of  $\lambda$  and  $c$ . For example he might

<sup>23</sup>This result is slightly different from the one obtained in Szalay (2005), as he was obtaining a smooth transition between the two regimes  $\Gamma_\lambda = \Gamma$  and  $\Gamma = \{0\}$ . This is because of the different assumptions on the information collection technology and the cost function associated to it.

want to induce some information partition that are cheap for the agent or try to repay him by delegating more than in the benchmark case. This might be a question for further research.

## 6 Conclusions

I study the optimal mechanism a principal should commit to in an environment where a biased agent can costly acquire information. The adverse selection in the reporting of information and the moral hazard due to the cost of information creates a non-monotonic effect of the cost of information on the utility of the principal.

# Appendix A

## Delegation

The optimal mechanism can be seen as a form of delegation to the agent with some form of discretion.

A (deterministic) mechanism with no transfers is a set of message  $M$  and an action  $x : M \rightarrow [-\frac{1}{2}, \frac{1}{2}]$  that specifies to which action the principal commits as a function of the message  $m \in M$ . A discretionary-delegation rule is a set  $\hat{\Gamma} \subseteq [-\frac{1}{2}, \frac{1}{2}]$ .

Given  $\hat{\Gamma}$  I can define a mechanism  $\{M, x\}$  that is equivalent to  $\hat{\Gamma}$ . Just take  $M = \hat{\Gamma}$  and  $x(m) = m$  for any  $m \in M$ . Starting from  $\{M, x\}$  I can define  $\hat{\Gamma} = x(M)$ . The two objects share the same properties.

### Proposition 1

I want to prove that the optimal contract consists of an interval. By contradiction I assume that it is not, or equivalently that there exist two actions  $a_1$  and  $a_2$  such that no other action  $a_1 < a < a_2$  is offered to the agent as part of the contract. It is possible to show that a profitable deviation exists. In fact if the principal offers to the agent the possibility to play  $a^* = \frac{a_1 + a_2}{2}$  she will see her expected payoff increase. In order to show this I will compute the expected payoff of the principal with and without  $a^*$  in the contract and compare them. When only  $a_1$  and  $a_2$  are offered to the agent he will buy a partition  $\psi = \{[\theta_0, \theta_1], \dots, (\theta_{i-1}, \theta_i], (\theta_i, \theta_{i+1}], \dots, (\theta_{n-1}, \theta_n], \text{ with } \theta_i = \frac{a_1 + a_2 - 2b}{2}\}$ <sup>24</sup> as in the following figure.

He will play  $a_1$  when  $\theta \in (\theta_{i-1}, \theta_i]$  and  $a_2$  when  $\theta \in (\theta_i, \theta_{i+1}]$ . The expected utility of the principal if the state is in  $[a_1 - b, a_2 - b]$  is given by:<sup>25</sup>

$$U = - \int_{a_1 - b}^{\frac{a_1 + a_2 - 2b}{2}} (a_1 - \theta)^2 d\theta - \int_{\frac{a_1 + a_2 - 2b}{2}}^{a_2 - b} (a_2 - \theta)^2 d\theta$$

When also  $a^*$  is offered

$$\psi = \left\{ [\theta_0, \theta_1^*], \dots, (\theta_{j-1}^*, \theta_j^*], (\theta_j^*, \theta_{j+1}^*], (\theta_{j+1}^*, \theta_{j+2}^*), \dots, (\theta_{m-1}^*, \theta_m] \right\}, \text{ with } \theta_j^* = \frac{a_1 + a^* - 2b}{2}, \theta_{j+1}^* = \frac{a_2 + a^* - 2b}{2}$$

I can compute the utility of the principal in the interval  $[a_1 - b, a_2 - b]$ :

<sup>24</sup>Remember that we are considering the case of low cost of information. Therefore the return on information is bigger than the cost. The agent chooses  $\theta_i$  so that his expected return on information is maximized.

<sup>25</sup>When the state is not in this interval the expected payoff of the principal is the same for the two contracts.

$$U^* = - \int_{a_1-b}^{\frac{a_1+a^*-2b}{2}} (a_1 - \theta)^2 d\theta - \int_{\frac{a_1+a^*-2b}{2}}^{\frac{a_1+a^*-2b}{2}} (a^* - \theta)^2 d\theta - \int_{\frac{a_2+a^*-2b}{2}}^{a_2-b} (a_2 - \theta)^2 d\theta$$

Simple algebra gives  $U^* - U \geq 0$ .

I want to find the largest action that the principal allows the agent to take. I first look for what the agent chooses to acquire when his menu of options<sup>26</sup> is  $[-\frac{1}{2}, \alpha]$ . The agent chooses  $\Gamma = \{\cup_{x \leq y} \{x\}, (y, \frac{1}{2})\}$ : below  $y$ , the agent learns the exact S.o.W. Above he knows that  $\theta \in (y, \frac{1}{2}]$  and plays  $\alpha$ .

$$\max_y - \int_{-\frac{1}{2}}^y (\theta + b - \theta - b)^2 d\theta - \int_y^{\frac{1}{2}} (\alpha - \theta - b)^2 d\theta - c \left( 1 - \left( \frac{1}{2} - y \right)^3 \right)$$

We can find a solution of the previous problem with the necessary and sufficient FOCs:  $(\alpha - y - b)^2 - 3c \left( \frac{1}{2} - y \right)^2 = 0$ , that is  $y = \frac{\alpha - b - \sqrt{3c}}{1 - \sqrt{3c}}$ .

The principal takes this value into consideration when she solves for the optimal threshold:

$$\begin{aligned} \max_{\alpha} - \int_{-\frac{1}{2}}^{y=\frac{\alpha-b-\sqrt{3c}}{1-\sqrt{3c}}} (\theta + b - \theta)^2 d\theta - \int_{y=\frac{\alpha-b-\sqrt{3c}}{1-\sqrt{3c}}}^{\frac{1}{2}} (\alpha - \theta)^2 d\theta \\ = -b^2 \left( y + \frac{1}{2} \right) + \frac{1}{3} [\alpha - \theta]_{y, \frac{1}{2}}^3 \end{aligned}$$

The F.O.C.s<sup>27</sup> give

$$\begin{aligned} F &= -b^2 \frac{dy}{d\alpha} + \left[ \frac{1}{2} - \alpha \right]^2 - [\alpha - y]^2 + [\alpha - y]^2 \frac{dy}{d\alpha} = \\ &= -\frac{b^2}{1-\sqrt{3c}} + \left[ \frac{1}{2} - \alpha \right]^2 + [\alpha - y]^2 \frac{\sqrt{3c}}{1-\sqrt{3c}} = \\ &= -\frac{b^2}{1-\sqrt{3c}} + \left[ \frac{1}{2} - \alpha \right]^2 + \frac{\sqrt{3c}}{(1-\sqrt{3c})^3} \left( 3c \left[ \frac{1}{2} - \alpha \right]^2 + b^2 + 2\sqrt{3c} \left[ \frac{1}{2} - \alpha \right] \right) = \\ &= -b^2 (1 + 3c - 3\sqrt{3c}) + \left[ \frac{1}{2} - \alpha \right]^2 6cb + \left[ \frac{1}{2} - \alpha \right]^2 (1 - 3\sqrt{3c} + 9c) = 0 \end{aligned}$$

First, note that, when  $c = \frac{1}{12}$  the previous expression is  $\frac{1}{4} \left[ \frac{1}{2} - \alpha \right]^2 + \frac{b}{2} \left[ \frac{1}{2} - \alpha \right] + \frac{1}{4} b^2 = 0$ , which gives as solution  $\alpha = \frac{1}{2} + b$ , that is not admissible. It must be  $\alpha = \frac{1}{2}$  when  $c$  is large enough ( $c = \hat{c}$ ). In order to check how the solution moves with the cost parameter  $c$  we use the IFT and compute:

<sup>26</sup>Note that in low SoW the principal cannot do better than in the case where  $c = 0$ , described in Melumad & Shibano (1991). Therefore delegation is optimal for low states.

<sup>27</sup>Sufficient

$$\frac{d(\frac{1}{2}-\alpha)}{dc} = -\frac{\frac{dF}{dc}}{\frac{d(\frac{1}{2}-\alpha)}{dc}} = -\frac{\frac{3\sqrt{3}}{2\sqrt{c}}(b^2 - [\frac{1}{2}-\alpha]^2) + [\frac{1}{2}-\alpha]^2 - 3b^2 + 6b[\frac{1}{2}-\alpha]}{2(\frac{1}{2}-\alpha)[1+9c-3\sqrt{3c}] + 6bc} = -\frac{\geq 0}{> 0} \text{ with } b^2 - [\frac{1}{2}-\alpha]^2 > 0 \text{ as } b > [\frac{1}{2}-\alpha] \text{ and } [\frac{1}{2}-\alpha]^2 - 3b^2 + 6b[\frac{1}{2}-\alpha] > 0.$$

Therefore  $[1-\alpha]$  decreasing in  $c$  and  $\alpha$  increasing in  $c$ .

### Lemma 1

Let's assume that the principal provide 3 possible options  $\{a_1, a_2, a_3\}$  to the agent. I want to show that it is not possible that the agent chooses an interval partition  $\{[-\frac{1}{2}, \theta_1], [\theta_1, \theta_2], [\theta_2, \frac{1}{2}]\}$  with  $-\frac{1}{2} < a_1 < \theta_1 < a_2 < \theta_2 < a_3 < \frac{1}{2}$ . WLOG assume  $a_2 < 0$ . The agent solves:

$$\max_{\theta_1, \theta_2} - \int_{-\frac{1}{2}}^{\theta_1} (a_1 - \theta)^2 - \int_{\theta_1}^{\theta_2} (a_2 - \theta)^2 - \int_{\theta_2}^{\frac{1}{2}} (a_3 - \theta)^2 - c \left( 1 - \left( \theta_1 + \frac{1}{2} \right)^3 - (\theta_2 - \theta_1)^3 - \left( \frac{1}{2} - \theta_2 \right)^3 \right)$$

The F.O.C.s give

$$\theta_1 : a_2^2 - a_1^2 - 2\theta_1(a_2 - a_1) + 3c[\frac{1}{4} - \theta_2^2 + \theta_1 + 2\theta_1\theta_2] = 0$$

$$\theta_2 : a_3^2 - a_2^2 - 2\theta_2(a_3 - a_2) + 3c[-\frac{1}{4} + \theta_1^2 + \theta_2 - 2\theta_1\theta_2] = 0$$

And the associated Hessian matrix is:

$$H = \begin{bmatrix} -2(a_2 - a_1) + (1 + 2\theta_2)3c & 6c(\theta_1 - \theta_2) \\ 6c(\theta_1 - \theta_2) & -2(a_3 - a_2) + (1 - 2\theta_1)3c \end{bmatrix}.$$

Given that  $b = 0$  if  $\theta_1, \theta_2$  is a maximum for the agent for some  $\{a_1, a_2, a_3\}$  then  $\theta_1, \theta_2$  would be a maximum when  $a_1 = \frac{-\frac{1}{2} + \theta_1}{2}$ ,  $a_2 = \frac{\theta_1 + \theta_2}{2}$ ,  $a_3 = \frac{\theta_2 + \frac{1}{2}}{2}$ . In fact, the agent and the principal wants the same action to be played (action=expectation of the posterior).

$$H = \begin{bmatrix} -(\frac{1}{2} + \theta_2) + (1 + 2\theta_2)3c & 6c(\theta_1 - \theta_2) \\ 6c(\theta_1 - \theta_2) & -(\frac{1}{2} - \theta_1) + (1 - 2\theta_1)3c \end{bmatrix}$$

$$\text{and in } c \rightarrow \frac{1}{12}^+ \quad H = \begin{bmatrix} -\frac{1}{4} - \frac{\theta_2}{2} & \frac{1}{2}(\theta_1 - \theta_2) \\ \frac{1}{2}(\theta_1 - \theta_2) & -\frac{1}{4} + \frac{\theta_1}{2} \end{bmatrix}.$$

The matrix is negative definite if  $a_2 = \frac{\theta_1 + \theta_2}{2} \leq 0$ , in fact  $-\frac{1}{4} - \frac{\theta_2}{2} < 0$  and  $\det(H) > 0$  for what follows:

$$\det(H) > 0 \text{ iff } 1 - 2\theta_1 + 2\theta_2 + 4\theta_1\theta_2 - 4\theta_1^2 - 4\theta_2^2 > 0$$

This condition is always satisfied unless  $\theta_1 = -\frac{1}{2}, \theta_2 = \frac{1}{2}$ .



This implies that if  $a_2 \leq 0$  there is only one solution for the problem of the agent. This solution is given by  $a_1 = -\frac{1}{3}$ ,  $a_2 = 0$ ,  $a_3 = \frac{1}{3}$ , with  $\theta_1 = -\frac{1}{6}$ ,  $\theta_2 = \frac{1}{6}$ . We notice now that this cannot be a solution because the agent would prefer not to buy any information and play  $a_2 = 0$ .

This implies that  $a_2 > 0$ , which contradicts the initial Hp. Hence it must be the case that the agent will never buy an information partition with more than 2 elements.

When  $b > 0$  I can rewrite the F.O.C.s with  $a'_1 = a_1 - b$ ,  $a'_2 = a_2 - b$  and  $a'_3 = a_3 - b$ . If I substitute in the new Hessian matrix  $a'_1 = \frac{-\frac{1}{2} + \theta_1}{2}$ ,  $a'_2 = \frac{\theta_1 + \theta_2}{2}$ ,  $a'_3 = \frac{\theta_2 + \frac{1}{2}}{2}$ , I obtain the same condition as in the case where  $b = 0$ .

### Proposition 2

When  $c < \frac{1}{6}$

$$\{0\} = \arg \max_{\theta'} - \int_{-\frac{1}{2}}^{\theta'} \left(-\frac{1}{4} - \theta\right)^2 d\theta - \int_{\theta'}^{\frac{1}{2}} \left(\frac{1}{4} - \theta\right)^2 d\theta - c \left(1 - \left(\theta' + \frac{1}{2}\right)^3 - \left(\frac{1}{2} - \theta'\right)^3\right)$$

When  $c > \frac{1}{6}$  I can find  $y$  as the solution of

$$- \int_{-\frac{1}{2}}^{\theta'} (-y - \theta)^2 d\theta - \int_{\theta'}^{\frac{1}{2}} (y - \theta)^2 d\theta - c \left(1 - \left(\frac{1}{2}\right)^3 - \left(\frac{1}{2}\right)^3\right) = - \int_{-\frac{1}{2}}^{\frac{1}{2}} (y - \theta)^2 d\theta - c (1 - (1)^3)$$

and then verify that

$$\{0\} \in \arg \max_{\theta'} - \int_{-\frac{1}{2}}^{\theta'} (-y - \theta)^2 d\theta - \int_{\theta'}^{\frac{1}{2}} (y - \theta)^2 d\theta - c \left(1 - \left(\theta' + \frac{1}{2}\right)^3 - \left(\frac{1}{2} - \theta'\right)^3\right)$$

### Proposition 3

We want to characterize the solution for the problem when  $c > \frac{1}{12}$  and  $b > 0$ . The principal solves:

$$\max_{a_1, a_2} EU(a_1, a_2) = - \int_{-\frac{1}{2}}^{\theta} (a_1 - \theta)^2 d\theta - \int_{\theta}^{\frac{1}{2}} (a_2 - \theta)^2 d\theta \text{ s.t.}$$

$$\theta \in \arg \max_{\theta'} EV(\theta') = - \int_{-\frac{1}{2}}^{\theta'} (a_1 - \theta - b)^2 d\theta - \int_{\theta'}^{\frac{1}{2}} (a_2 - \theta)^2 d\theta - c \left(1 - \left(\theta' + \frac{1}{2}\right)^3 - \left(\frac{1}{2} - \theta'\right)^3\right)$$

In order to make computation easier I adopt the following change of variables  $\begin{cases} x = a_1 + a_2 \\ y = a_2 - a_1 \end{cases}$  with

$$\begin{cases} x \in [0, 1 - \|y\|] \\ y \in [-1, 1] \end{cases}$$

We have  $EU = -\frac{1}{12} + \left(\frac{1}{4} - \theta^2\right)y - \frac{1}{4}(x^2 + y^2) + \theta xy$  and  $EV(a_1, a_2) = EU(a_1, a_2) - b^2 + bx - 2\theta by - \frac{3}{4}c + 3c\theta^2$

The constraint of the the agent gives  $\frac{d}{d\theta}EV = 0 \Rightarrow g = 6c\theta - 2\theta y + xy - by = 0$

$$\theta = \frac{1}{2} \frac{y(x - 2b)}{y - 3c}$$

with  $\frac{d^2}{d\theta^2}EV \leq 0$ , *iff*  $y - 3c \geq 0$ . This implies that we have only one solution when

$$y - 3c > 0$$

Not that when  $c \rightarrow \frac{1}{12}^+$  the previous condition reads  $y > \frac{1}{4}$ . This would always be the case, otherwise the contract  $a_1 = -\frac{1}{4} + b$ ,  $a_2 = \frac{1}{4} + b$  would do better for the principal while not solving the FOCs. I will analyze what happens when  $c$  is large in the next corollary.

We have	$\frac{\delta U}{\delta x} = -\frac{1}{2}x + \theta y$	$\frac{\delta g}{\delta x} = y$
	$\frac{\delta U}{\delta y} = \frac{1}{4} - \theta^2 - \frac{1}{2}y + \theta x$	$\frac{\delta g}{\delta y} = x - 2\theta - 2b$
	$\frac{\delta U}{\delta \theta} = -2\theta y + xy$	$\frac{\delta g}{\delta \theta} = -2y + 6c$

The F.O.C.s for the principal give:

$$\frac{\delta U}{\delta x} + \lambda \frac{\delta g}{\delta x} = 0, \frac{\delta U}{\delta y} + \lambda \frac{\delta g}{\delta y} = 0, \frac{\delta U}{\delta \theta} + \lambda \frac{\delta g}{\delta \theta} = 0, g = 0.$$

From which:

$$\frac{\delta U}{\delta \theta} + \lambda \frac{\delta g}{\delta \theta} = 0 \Rightarrow \lambda = \frac{x(1-y)}{6c}$$

Note that we can also compute

$$\frac{dU}{db} = \frac{\delta U}{\delta b} + \lambda \frac{\delta g}{\delta b} = 0 + \frac{x(1-y)}{6c} (-2y) = -\frac{xy(1-y)}{3c}$$

from which we get that  $x \geq 0$  in order to have  $\frac{dU}{db} \leq 0$

---

The FOC on  $x$  gives

$$x : \frac{dU}{dx} = 0 \Rightarrow \frac{1}{2} \left[ -x + 2\theta y + (xy - 2\theta y) \frac{y}{y-3c} \right] = 0 \Rightarrow x = \frac{6c\theta y}{(y^2 - y + 3c)}$$

From  $x \geq 0$  we have  $\theta \geq 0$  (note that  $y^2 - y + 3c > 0$  for  $c > \frac{1}{12}$ ).

This implies also that

$$x > 2b$$

otherwise we would have  $\theta < 0$  from the optimization problem of the agent.

The FOC on  $y$  gives

$$y : \frac{dU}{dy} = 0 \Rightarrow \frac{1}{4} - \theta^2 - \frac{1}{2}y + \theta x + (x - 2\theta) y \frac{1}{2} \frac{x - 2\theta - 2b}{y - 3c} = 0 \Rightarrow$$

$$\frac{1}{4} - \theta^2 - \frac{1}{2}y + \theta x + \frac{2\theta - x}{y - 3c} 3c\theta = 0 \text{ which gives (substituting } x = \frac{6c\theta y}{(y^2 - y + 3c)})$$

$$\begin{aligned} & \frac{1}{4} - \frac{1}{2}y + \frac{\theta^2}{(y^2 - y + 3c)(y - 3c)} \left[ -(y - 9c)(y^2 - y + 3c) + 6cy(y - 6c) \right] = \\ & = \frac{1}{2} - y + \frac{2\theta^2}{(y^2 - y + 3c)} \left[ -y^2 + y(12c + 1) - 9c \right] = 0 \end{aligned}$$

Note that  $y \neq \frac{1}{2}$  because otherwise the FOC on  $x$  would not be satisfied.

If, by contradiction, it was  $y > \frac{1}{2} \Rightarrow -y^2 + y(12c + 1) - 9c > 0$  which means

$$y > \frac{1}{2} + \frac{1}{6}c - \frac{\sqrt{(12c-1)^2 + 12c}}{2}.$$

When  $c$  is large enough this would imply that the principal loses (for example  $c = \frac{1}{9}$   $y > 0.56$ ) whatever is  $a_1, a_2$  and  $\theta$ , even  $\theta = 0$ ,  $a_1 = -\frac{0.56}{2}$ ,  $a_2 = \frac{0.56}{2}$  she would get better by choosing  $a_1 = -\frac{1}{4} + b$ ,  $a_2 = \frac{1}{4} + b$  inducing  $\theta = 0$ .<sup>28</sup>

Remind that it cannot be  $y = \frac{1}{2}$  and  $y$  should be continuous in  $c$ .

Hence it must be

$$y < \frac{1}{2}.$$

### Corollary 1

From the previous proposition we know that in order to have  $\theta$  being a solution of the maximization problem of the agent it must be  $y - 3c > 0$ . When  $c \geq \frac{1}{6}$  this condition is  $y \geq \frac{1}{2}$  meaning that the FOCs found in proposition 3 do not have a solution.

---

<sup>28</sup>if  $b$  is small enough

We will have that  $y = 3c$  and the agent will be indifferent among any  $\theta$  (any form of information acquisition). Hence, for some  $\frac{1}{12} < c^* \leq \frac{1}{6}$  and for any  $c > c^*$  we have  $y = 3c$  and  $x = 2b$ .

$\theta$  <sup>29</sup> maximizes  $U = -\frac{1}{12} + \left(\frac{1}{4} - \theta^2\right) 3c - \frac{1}{4} (9c^2 + 4b^2) + \theta 3c 2b$  giving

$$\theta = b.$$

#### Proposition 4

From Proposition 1 we can compute  $\frac{dU}{dc}$  for  $c \in (0, \frac{1}{12})$  and observe that it is negative.

Proposition 3 gives an expression for  $\frac{dU}{dc}$  in  $c \in (0, c^*)$

$$\frac{dU}{dc} = \frac{\delta U}{\delta c} + \lambda \frac{\delta g}{\delta c} = 0 + \frac{x(1-y)}{6c} (6c\theta) = x(1-y)\theta > 0$$

When  $c > c^*$  we have  $y = 3c$  and

$$EU = -\frac{1}{12} + \frac{3}{4}c + 3b^2c - \frac{9}{4}c^2 - b^2$$

Hence  $\frac{dEU}{dc} = \frac{3}{4} + 3b^2 - \frac{18}{4}c$  that is  $\geq (<) 0$  iff  $c \leq (>) \frac{1}{6} + b^2$ .

#### Proposition 5

First, note that we can think of this problem as one where the principal maximizes the expected value of both parties (weighted by  $\lambda$  and  $(1 - \lambda)$ ) and she bears the cost of the information collection.

When  $c \leq \frac{1}{12}$  we know that the agent is better off by buying as much information as possible. Therefore we just need to verify for what values of  $\lambda$  the principal wants to induce full information acquisition or no information at all. When the agent buys “full information” he gets  $V = (1 - \lambda)[0 - c]$  and the principal gets  $U = \lambda(0)$ , we have to verify when  $V + U > -\lambda\frac{1}{12}$ , that is the utility the principal gets in absence of information and without the need to make transfers to the agent. The condition is satisfied iff

$$\lambda > \frac{c}{c + \frac{1}{12}}$$

Note that RHS is increasing in  $c$  and takes value  $\frac{1}{2}$  when  $c = \frac{1}{12}$  as we could expect.

When  $c \in (\frac{1}{12}, \frac{1}{3}]$  we know from Lemma 1 that a maximum of two actions can be induced. The principal chooses the information partition (compatible with the incentive of the agent) that maximizes  $V + U$  and

---

<sup>29</sup>conveniently for the principal

compare this with no information acquisition.

When  $c \in (\frac{1}{12}, \frac{1}{6}]$

$$\max_{\theta'} - \int_{-\frac{1}{2}}^{\theta'} \left( \frac{-\frac{1}{2} + \theta'}{2} - \theta \right)^2 d\theta - \int_{\theta'}^{\frac{1}{2}} \left( \frac{\frac{1}{2} + \theta'}{2} - \theta \right)^2 d\theta - (1 - \lambda) c \left( 1 - \left( \theta' - \frac{1}{2} \right)^3 - \left( \frac{1}{2} - \theta' \right)^3 \right) \geq (<) - \lambda \frac{1}{12}$$

$$s.to \ \theta' \in \arg \max_{\hat{\theta}} - \int_{-\frac{1}{2}}^{\hat{\theta}} \left( \frac{-\frac{1}{2} + \hat{\theta}}{2} - \theta \right)^2 d\theta - \int_{\hat{\theta}}^{\frac{1}{2}} \left( \frac{\frac{1}{2} + \hat{\theta}}{2} - \theta \right)^2 d\theta - c \left( 1 - \left( \hat{\theta} - \frac{1}{2} \right)^3 - \left( \frac{1}{2} - \hat{\theta} \right)^3 \right)$$

The principal will choose  $\theta' = 0$  and compare the total utility the two agents get and find that the previous condition is satisfied for

$$\lambda > \frac{1 + 36c}{4 + 36c},$$

increasing in  $c$ .

Similarly, when  $c \in (\frac{1}{6}, \frac{1}{3})$  the principal will solve

$$-2 \int_{-\frac{1}{2}}^0 \left( \frac{3}{2}c - \theta \right)^2 d\theta - (1 - \lambda) c \left( 1 - \left( \frac{1}{2} \right)^3 - \left( \frac{1}{2} \right)^3 \right) \geq (<) - \lambda \frac{1}{12}$$

and find that this condition is satisfied for

$$\lambda > \frac{\frac{1}{12} + \frac{9}{4}c^2}{\frac{1}{12} + \frac{3}{4}c},$$

that is increasing in  $c$  and  $\lim_{c \rightarrow \frac{1}{3}} \frac{\frac{1}{12} + \frac{9}{4}c^2}{\frac{1}{12} + \frac{3}{4}c} = 1$ .

## Appendix B

In this section I want to further analyze the role of the cost functions on the results. Consider two information partition  $\psi = \psi_A$  and  $\psi'' = \psi_B$  as in the Figure 1. The assumption on the cost function implies that  $c(\psi_A) > c(\psi_B)$ .

**Remark 1**

Theorem 1 of Blackwell (1953) tells us that there is a sense for which, under any reasonable measure of uncertainty, the experiment that gives partition  $\psi_A$  is more precise than the one from partition  $\psi_B$ . My choice for the cost function corresponds to  $\phi(\cdot) = (\cdot)^2$  in Theorem 1 of Blackwell (1953). Any other convex function would have been coherent with the theory elaborated in Blackwell (1953), but would in any case have the same implication on how to order  $\psi_A$  and  $\psi_B$  in terms of informativeness.

**Remark 2**

What is important to stress is that the choice of the cost function buys me:

1. the full characterization of the optimal contract
2. A clear switch of regime from infinite actions chosen in the optimal contract with  $c \leq \frac{1}{12}$  to only 2 actions when  $c > \frac{1}{12}$ .

Whether 1. can be seen as a clear advantage for tractability, 2. might be a weakness and a possible way for new research.

A further implication due to the choice in the cost function is that it is possible to rank and compare the cost of information with respect to its value.

**Remark 3**

One possible way to go away from the original Hp on the cost function is to assume that the cost of each partition distribution is given by the number of cuts that the agent decides to do, so that the partition  $\psi_A$  and  $\psi_B$  have the same cost. In order to compare this case with the one studied before I will compute which 2 actions the principal optimally decides to offer to the agent. I look for the solution of the problem:

$$\begin{aligned} & \max_{a_1, a_2} - \int_{-\frac{1}{2}}^{\theta'} (a_1 - \theta)^2 d\theta - \int_{\theta'}^{\frac{1}{2}} (a_2 - \theta)^2 d\theta \\ & \text{s.to } \theta' \in \arg \max_{\theta} - \int_{-\frac{1}{2}}^{\theta} (a_1 - \theta - b)^2 d\theta - \int_{\theta}^{\frac{1}{2}} (a_2 - \theta - b)^2 d\theta \end{aligned}$$

This problem has as solution  $\theta' = \frac{a_1 + a_2}{2} - b$ ,  $a_1 = -a_2$  and  $a_1 + a_2 = \frac{1}{2} - b^2$ .

As in the case developed in the paper the condition  $a_1 + a_2 = \frac{1}{2} - b^2 < \frac{1}{2}$  tells us that the principal is ready to lose something in terms of informativeness in the induced information acquisition in order to reduce the adverse selection.

### Different technology

I want to show what happens when I do not assume that the technology to learn information is a different one. I assume that the agent can exert low or high effort  $0 < e_L < e_H = 1$  at cost  $c_L < c_H$  and learn the state with probability  $e_L$  and  $e_H$ , respectively. While it is an interesting question to try to solve for the optimal mechanism, here I will only compare which of two contract,  $\Gamma^1 = [-\frac{1}{2}, \frac{1}{2}]$  or  $\Gamma^2 = [-\frac{1}{2}, \frac{1}{2} - k]$  the planner should use. If the contract chosen by the principal is  $\Gamma^1$  the agent choosing  $e_L$  or  $e_H$  has expected utility respectively

$$v_L^1 = -(1 - e_L) \int_{-\frac{1}{2}}^{\frac{1}{2}} (b - (\theta + b))^2 d\theta - e_L \int_{\frac{1}{2}-b}^{\frac{1}{2}} \left( \frac{1}{2} - (\theta + b) \right)^2 - c_L$$

$$v_H^1 = - \int_{\frac{1}{2}-b}^{\frac{1}{2}} \left( \frac{1}{2} - (\theta + b) \right)^2 - c_H$$

Similarly we can find the utility of the agent when the contract offered is  $\Gamma^2$  and note that  $v_L^1 > v_L^2, v_H^1 > v_H^2$ . If  $c_L$  is much smaller than  $c_H$  the agent would typically have  $v_L^2 > v_H^2$  when  $k$  is large, but he might prefer  $v_H^1$  to  $v_L^1$ . The principal's utility will be larger when the agent exert more effort pushing her to choose the contract  $\Gamma^1$ , choosing therefore to delegate more.

More delegation as it was happening in the case of low cost information in Proposition 1 could incentivize information collection and therefore benefits the principal.

## References

- [1] Alonso R. , Matouschek N. (2007): “Relational Delegation”, *RAND J. Econ.* 38 (4), 1070–1089.
- [2] Angelucci C. (2017): “Motivating Agents to Acquire Information”, WP.
- [3] Brown R.L. (2010): “Measuring Delegation” *Rev Int Organ* (2010) 5:141–175.
- [4] Argenziano R., Severinov S., Squintani F. (2016): “Strategic Information Acquisition and Transmission”, *American Economic Journal: Microeconomics*, 8(3): 119-155.
- [5] Blackwell, D. (1953). “Equivalent comparisons of experiments.”, *The Annals of Mathematical Statistics*. 24(2): 265-272.
- [6] Cai H. (2009): “Costly participation and heterogeneous preferences in informational committees”, *The RAND Journal of Economics*, 40(1), 173-189.
- [7] Clinton J.D., Bertelli A., Grose C.R., Lewis D.E., Nixon D.C. (2011): “Separated Powers in the United States: The Ideology of Agencies, Presidents, and Congress”, *AJPS*, Volume56, Issue2 pp 341-354
- [8] Crawford V. , Sobel J. (1982): “Strategic information transmission”, *Econometrica* 50 (6) 1431–1451.
- [9] Deimen I, Szalay D. (2017): “Delegated expertise, authority, and communication” R&R AER.
- [10] Dessein W. (2002): “Authority and Communication in Organizations” *Review of Economic Studies* (2002).
- [11] Dewatripont M., Tirole J. (1999), “Advocates”, *The Journal of Political Economy*, Vol. 107, No. 1. (Feb., 1999), pp. 1-39.
- [12] Di Pei H.(2015): “Communication with Endogenous Information Acquisition”: *Journal of Economic Theory*, Volume 160.
- [13] Gailmard S.(2009): “Discretion Rather than Rules: Choice of Instruments to Control Bureaucratic Policy Making”, *Political Analysis*,17:25–44
- [14] Gentzkow M., Kamenica E. (2011): “Bayesian Persuasion” *American Economic Review* Vol. 101(6), pp. 2590-2615



- [15] Gentzkow M., Kamenica E. (2014): "Costly Persuasion" American Economic Review: Papers & Proceedings Vol. 104(5), pp. 457-462.
- [16] Gerardi D., Yariv L. (2008): "Information Acquisition in Committees", Games and Economic Behavior, 62, 436-459.
- [17] Gershkova A, Szentesi B. (2009): "Optimal voting schemes with costly information acquisition". JET, 144(1), 36-68.
- [18] Goltsman M., Horner J., Pavlov G., Squintani F.(2009): "Mediation, Arbitration and Negotiation", Journal of Economic Theory, 144(4): 1397-1420.
- [19] Green J., Stokey, N. (1981), "The Value of Information in the Delegation Problem", Discussion Paper Harvard University.
- [20] Holmstrom B. (1978): "On Incentives and Control in Organization" Phd Dissertation Stanford University.
- [21] Holmstrom B. (1979): "Moral Hazard and Observability" Bell Journal of economics 10, 74-91.
- [22] Li H. (2001), "A Theory of Conservatism," Journal of Political Economy, 109(3), 617-636.
- [23] Melumad N., Shibano T. (1991): "Communication in Setting with No Transfers", Rand Journal of Economics 22(2), 173-198.
- [24] Selin J. L. (2015): "What Makes an Agency Independent?", AJPS, Volume59, Issue4, pp 971-987.
- [25] Sobel, J. (2013): "Giving and receiving advice." Advances in Economics and Econometrics.
- [26] Stephenson M. (2011): "Information Acquisition and Institutional Design", 124 Harv. L. Rev. 1422.
- [27] Szalay D. (2005): "The Economics of Extreme Options and Clear Advice": Review of Economic Studies 72, 1173-1198.

# Chapter 2

## Optimal intervention for network games

Julie Pinole, Lorenzo Verstraeten

### 1 Introduction

We study a classical network game. Players' payoffs are a function of their own characteristics, their actions and the actions of their peers. We ask how a social planner should intervene on the connections among players when she has in mind a particular objective. The planner might want to implement a given outcome, in that case we show how she can achieve this goal at minimal cost; or she might have some resources to spend, in that case we ask how she should allocate them in order to maximize aggregate welfare.

The type of intervention we have in mind is one that change the intensity of the connections among players. Some examples of interventions could be the increase (or decrease) of the number of extracurricular activities provided in a school or neighborhood context: adding book clubs, sport clubs, or any types of groups, both physical and virtual. The social planner can also play on the cost of those activities to incentivize or disincentivize participation. This will likely create or destroy links between individuals. We can think of a regulator influencing the partnership structure of firms, by encouraging local links between firms of different size or different centrality in a national or global scale versus promoting networks of firms similar in terms of characteristics. For instance the regulator may have voice to the chapter when financial cross-participation of big firms are realized. When agents are individuals, the planner can devise policies promoting the integration of newly arrived immigrants. The principal of a school could organize study groups where she decides their composition. For example, Algan et al. (2015) ([?]) ran an experiment in a French university by randomly assigning students into first-year groups. This design allowed to measure both the actual change in the network structure and whether it affected the outcome of interest.<sup>30</sup>

We assume that the intervention of the planner has increasing marginal cost. The more she wants to change a connection between two individuals the larger is the marginal cost she has to pay. We first ask how she could achieve, at minimal cost, a specific outcome for the network game. Modifying the connection

---

<sup>30</sup>This paper validates our assumption that in some circumstances a social planner can actually affect a social network.

between two individuals changes their incentives. This affects the decisions of all their peers and therefore the equilibrium of the network game. When the planner has in mind a specific equilibrium to implement we describe how she should modify the network structure in order to obtain her goal at minimal cost. The first set of results in Section 4 describes the interventions the planner should take and the cost she will incur. We manage to be very general in the results we obtain. In fact, we provide results for directed and undirected network. We also describe what happens when we allow the planner to change only a specific set of connections. While the mathematics behind all these results is the one developed in the computer science literature to solve the "nearest matrix" problem<sup>31</sup> its application to the study of network games is completely new. At the best of our knowledge we are the first to think of a minimal cost intervention by the planner as a minimal change in the adjacency matrix representing the network.

While the results obtained in section 4 are interesting per se, we could build on them to address some other questions. We ask how a social planner, with limited resources, can modify interactions among players in order to maximize aggregate welfare. Galeotti, Golub, Goyal (2018) ([?]), GGG from now on, are interested to the same problem. However, they focus on how a social planner should change players' incentives with an intervention targeting individual private values. The type of intervention we consider is instead one where the social planner intervenes on the network structure. Following GGG, we consider the effect of the policy on two kind of network games. The first category are games of strategic complements; in this case players are incentivized to engage in an activity if their peers do. The second category are games of strategic substitute; in this case players' incentives to engage in an activity are smaller the higher is the involvement of their neighbors. As in GGG, we draw conclusions on how qualitatively different the intervention is depending on whether we play a game of strategic complements or strategic substitutes.

We first try to understand which players will be more affected by the intervention of the planner. We decompose the effect of the intervention on a particular system of coordinates. The orthonormal basis obtained by diagonalizing the matrix representing the interactions between individuals. We show that the equilibrium of the network game can be measured in terms of the singular vectors of the adjacency matrix of the initial network. This is interesting because it allows us to understand how the planner change the players' incentives. We show that our intervention shares common features with the characteristics-intervention problem of GGG: in game of strategic complements central players will be mostly affected by the intervention; in game of strategic substitutes, instead, the incentives of neighbors are moved in opposite

---

<sup>31</sup>See Higham (2000)

directions. While the similarity with GGG for this result seems reasonable, it was not at all obvious ex ante; in fact the type of policy considered are quite different. Conceptually: in GGG the planner is limited to change only the private benefit of the actions of the players, therefore it is impossible for her to modify the influence of one player on another one. Mathematically: in GGG the planner moves an  $n$ -dimensional object (the vector of private benefit) while we give her the possibility to change a  $n \times n$  dimensional object (the full adjacency matrix). Even if the incentives of the players are moved similarly in our case as in GGG the way the social planner obtains her result using the two policies will be different.

In section 7 we try to analyze how the social planner affects the network structure. In the complement case, we show that the decision of the planners depends on two aspects. It is important whether players are central or not, and whether they have a high private marginal benefit for the action or not. If central players tend to have high private value, then the effort of the planner is unambiguously directed at them. The planner will sponsor links from and towards these players. Otherwise the result would depend on which aspect is the most important. The substitute case is more delicate. The planner will try to eliminate links in order to form a bipartite network.<sup>32</sup> She will try to form two groups in the population. Individuals inside one group share few links across them, while most of the links are across individuals of the two different groups. She achieves this by destroying links of low intensity.

The paper is organized as follows. Section 2 compares our work with the related literature. Section 3 describes the model. In section 4, starting from a given network structure, we characterize the minimal cost intervention to implement a given outcome. Section 5 exposes the planner's problem and gives properties of the equilibrium profile played on the new network. Section 6 provides a comparison between those properties and the results of GGG. Section 7 provides an analysis of the changes that the network structure goes through. Section 8 concludes.

## 2 Literature Review

We study a model of game on networks that covers many important situations. They fall into two main categories, as described by Bramoullé and Kranton (2016) ([?]): peer effects (games of strategic complements) and local public goods (games of strategic substitutes)<sup>33</sup>. Examples of outcomes where peer effects

<sup>32</sup>In graph theory, a bipartite graph is a graph whose vertices can be divided into two disjoint and independent sets  $U$  and  $V$  such that every edge connects a vertex in  $U$  to one in  $V$ .

<sup>33</sup>we abstract from the third topic they address, technology adoption, as adapting our model to discrete decisions is left for future work)

play a role range from smoking (Robalino and Macy, 2018, [?]), obesity (Trogdon et al., 2008, [?]), school achievement (Boucher et al., 2014, [?]), delinquent behaviors (Glaeser et al., 1996, [?]), to retirement savings (Saez and Duflo, 2003, [?]).<sup>34</sup> In all those situations, I am more likely to engage in an activity if my peers do. Those games are called games of strategic complementarities. On the contrary local public goods games exhibit strategic substitutability. I am less likely to contribute to a non-excludable good if my peers do and I can benefit from their contributions at zero cost. Bramoullé and Kranton, 2007 ([?]) depicts various interactions of this type. For instance information and innovation are often non-excludable. If my friends engage in information acquisition on a new consumption good, I may take advantage of it. Research and development expenses in enterprise generate innovations that also profit connected partners. Another dimension of technological spillovers is geography, as evidenced by Bloom et al., 2013, ([?]). A last example is what the literature refers to as crime games. Ballester et al., 2010 ([?]) quote the criminology literature to support their assumption that criminal skills are mostly learnt through peers, and thus there is spillover of crime activities from one individual to his connections. In all those examples, the exact structure of the network matters and the aggregate outcome is relevant to policy makers.

Our work represents a new application of the computer science literature on nearest matrices. Higham, 2000 ([?]) proposes answers to different mathematical problems searching for a matrix with specific properties that is as close as possible to an initial matrix deprived of this property, where different closeness metrics are possible. We provide a specific application of the nearest matrix problem by defining the network structure through its adjacency matrix and interpreting the nearest matrix solving an adequate optimization problem as the adjacency matrix of the desired network structure.

Finally, our work contributes to the literature on optimal strategy in the presence of social interactions. Zenou, 2016 ([?]) provides, among other things, a review of the literature on network intervention in games. Among the economics literature we quote other recent works: Fainmesser and Galeotti, 2016 ([?]), Akbarpour, Malladi, and Saberi, 2017 ([?]), Banerjee, Chandrashekar, Duflo, and Jackson, 2016 ([?]), Candogan, Bimpikis, and Ozdaglar, 2012 ([?]). Other disciplines investigate the topic. In marketing and computer science, the problem is often whom to target: Borgatti, 2006 ([?]), Kempe, Kleinberg, and Tardos, 2003 ([?]).

---

<sup>34</sup>Bramoullé et al., 2009, [?], and Boucher and Fortin, 2016, [?] provide interesting studies of peer effects with a focus on the associated econometrics challenges.

## 3 The Model

### 3.1 The setup

We study a game where  $n$  players are located on a directed network described by the weighted adjacency matrix  $G \in \mathcal{M}_{n,n}$ .<sup>35</sup> The set of player is called  $\mathcal{N} = 1, \dots, n$ . The element  $g_{ij}$  of  $G$  represents how strong the connection between players  $i$  and  $j$  is. We impose  $g_{ij} \geq 0$  for all  $i, j$ .

Each player  $i$  chooses an action  $a_i$  from  $\mathbb{R}_+$ . We call  $a \in \mathbb{R}_+^n$  the vector containing the action profile of all the players:  $a = (a_i)_{i \in \mathcal{N}}$ .

The payoffs to individual  $i$  are given by:

$$W_i(\mathbf{a}) = b_i a_i - \frac{1}{2} a_i^2 + \beta a_i \sum_{j \in \mathcal{N}} g_{ij} a_j$$

where  $b_i \in \mathbb{R}_+^*$  is an individual-specific characteristic measuring individual  $i$ 's marginal return of the direct effect of his action. We call  $b$  the vector containing the characteristics of the  $n$  players. Each player incurs a quadratic cost. Finally, each player's payoffs are affected by the interaction between his own action and the action of his connections. If  $\beta$  is positive, actions are strategic complement, whereas if  $\beta$  is negative they are strategic substitutes.

For notation purposes we call  $A$  the following transformation of the network structure:

$$A \equiv I - \beta G \tag{3.1}$$

### 3.2 Equilibrium

We impose the following assumption:

**Assumption 2.** *The spectral radius of  $\beta G$  is less than 1.*

Under assumption (2), Bramoullé, Kranton and D'Amours (2014) [?] shows that there exists a unique equilibrium  $\mathbf{a}^*$  for the network game described above. Furthermore the equilibrium satisfies the following system of linear equations:

$$[I - \beta G] \mathbf{a}^* = b \tag{3.2}$$

---

<sup>35</sup>For some of our results we allow  $g_{ii}$  to be different from 0, with the interpretation that it is a factor that influence the cost of player  $i$  when he chooses action  $a_i$ . In terms of the network structure this is equivalent to assume the presence of self-loops.

Assumption (2) ensures that  $I - \beta G$  is invertible. Hence we can write

$$\mathbf{a}^* = [I - \beta G]^{-1} \mathbf{b}$$

## 4 Closest network structure to implement a chosen vector of actions

We assume that the social planner can intervene to change the structure of the network. She could do this by changing the intensity of links across players. In this section we ask how she should do it if there is a specific action profile  $\bar{\mathbf{a}}$  that she wants to implement at minimal cost. For now the goal of the social planner is to implement a given equilibrium. She might for example want to get an outcome where all the players contribute equally (they all play the same action in equilibrium) or she might want all the players to increase their action by a certain amount. We characterize here the minimal intervention she should design and the associated cost. We make the following assumption about the cost she incurs to alter the network:

**Assumption 3.** *Changing the structure of the network from  $G$  to  $G^*$  has a cost of  $\|G^* - G\|_F$ , where  $\|M\|_F = \sqrt{\sum_{ij} m_{ij}^2}$  is the Frobenius norm of  $M$ .*

The previous assumption captures the idea that the social planner can modify the interaction between  $i$  and  $j$  at a convex cost. The planner faces increasing marginal costs of intervention. The Frobenius norm can be seen as an extension of the euclidean norm to  $\mathcal{R}^{n \times n}$ . The closest the new adjacency matrix is to the initial one the smaller the cost for the planner.

In the computer science literature there exists a lot of works aimed at finding the closest matrix to a given one. Often when an algorithm is implemented some approximations are used and the matrix of interest might lose some properties. This is why we might want to substitute the matrix obtained through the algorithm with the closest one having the desired properties.<sup>36</sup> We will use the results in this literature with a completely different purpose. The idea is to rewrite the planner's problem as one of finding the closest matrix to a given one. At that point the mathematical challenges are identical to the one faced (for completely different reason) in that literature and we can therefore borrow several results from there.

**Definition 1.** *We call  $Q(y, x)$  the set of matrix quotients of  $y$  by  $x$ , with  $y, x \in \mathbb{R}^n$ :*

$$Q(y, x) = \{M \in \mathcal{A} | Mx = y\} \tag{4.1}$$

---

<sup>36</sup>see Higham (2000) [?].

where  $\mathcal{A} \subseteq \mathcal{M}_{n,n}$  is a set of matrices with some desirable properties.

**Definition 2.** We call  $\mu^{x,A}(y)$  the minimal cost of altering the matrix  $A$  so that the resulting matrix belongs to  $Q(y, x)$ :

$$\mu^{x,A}(y) = \min_{E \in \mathcal{M}_{n,n}} \{ \|E\|_F : (A + E) \in Q(y, x) \cap \mathcal{A} \} \quad (4.2)$$

where  $\mathcal{A} \subseteq \mathcal{M}_{n,n}$  is a set of matrices with some desirable properties.

Higham, 2000 ([?]) reviews results of the computer science literature that solves the minimization problem (4.2) for different constraints on the type of matrices to work with (constraints defined in the set  $\mathcal{A}$ ). This is of interest for the problem at hand as the equilibrium condition (3.2) makes our problem equivalent to the minimization problem (4.2) with  $A = I - \beta G$ ,  $x = b$  and  $y = \bar{a}$ .

We focus on the three types of constraints that we consider the most relevant: the unconstrained case, the case where the starting network is undirected and we wish to reach an undirected network as well, and the case where we want to preserve some sparseness properties of the network.

#### 4.1 Case 1: Unconstrained intervention

The first type of intervention we consider is one where we don't impose any constraint to the planner. A given initial network is given and she can change any connection at a cost specified in Assumption 2. We will provide a closed form solution to the problem of the planner and specifies the cost that is associated to the intervention

For simplicity, we are now using the notation  $A$  defined in equation (3.1). Not imposing any condition on the intervention of the planner translates in the language of the previous definitions in  $\mathcal{A} = \mathcal{M}_{n,n}$ . The following proposition gives us the result:

**Proposition 6.** *The least costly intervention such that the action profile  $\bar{a}$  is played in equilibrium in the game played on the transformed network is:*

$$A + E_{\min}(\bar{a}) \quad (4.3)$$

with



$$E_{\min}(\bar{\mathbf{a}}) = \frac{(\mathbf{b} - A\bar{\mathbf{a}})\bar{\mathbf{a}}^T}{\bar{\mathbf{a}}^T \bar{\mathbf{a}}} \quad (4.4)$$

and this closest matrix is reached at a cost of:

$$\mu^{b,A}(\bar{\mathbf{a}}) = \frac{\|\mathbf{b} - A\bar{\mathbf{a}}\|}{\|\bar{\mathbf{a}}\|} \quad (4.5)$$

Note that  $\|\cdot\|$  denotes the euclidean norm of  $\mathbb{R}^n$ .

*Proof.* We first quote the following lemma, which is a result coming from the section 8 of Higham, 2000 ([?]):

**Lemma 2.** *Given  $y, x \in \mathbb{R}^n$ ,  $A \in \mathcal{M}_{n,n}$ , the following minimization problem:*

$$\min_{E \in \mathcal{M}_{n,n}} \{\|E\|_F : (A + E) \in Q(x, x)\}$$

*admits  $E_{\min}$  as a solution with*

$$E_{\min} = \frac{(y - Ax)x^T}{x^T x}$$

*and reach  $\mu^{x,A}(y)$  as minimum with*

$$\mu^{x,A}(y) = \frac{\|y - Ax\|}{\|x\|}$$

This result tells us how to change a matrix in the sense of minimizing the Frobenius norm of the difference between the initial and the final matrix, subject to the matrix belonging to  $Q(y, x)$ .

The proof of Proposition (6) is a direct application of this lemma for  $y, x = \bar{\mathbf{a}}, \mathbf{b}$ .

The following corollary simply uses (3.1) to go from the transformed matrix of  $A$  to the transformed matrix of  $G$ :

**Corollary 1.** *The new network structure in the modified game is:*

$$G(\bar{\mathbf{a}}) = G - \frac{1}{\beta} E_{\min}(\bar{\mathbf{a}})$$

*reached at a cost of  $\frac{\|\mathbf{b} - (I - \beta G)\bar{\mathbf{a}}\|}{\|\bar{\mathbf{a}}\|}$ .*

**Remark 1.** Mathematically  $E_{\min}$  is what the planner adds to  $A$  to reach the transformed matrix. That is why we call  $E_{\min}$  the optimal intervention. The first thing to note is that the matrix  $E_{\min}$  is of rank 1. Therefore we say that the planner's optimal intervention is a rank-1 intervention, that is of low computational complexity. This is true independently of the initial conditions. The second thing to note is that the optimal intervention  $E_{\min}$  is a priori non symmetric; this is the case even when the initial matrix is symmetric.

With this type of intervention, we are not constraining the social planner in any way. This is something that might not be desirable in some situations. When we start from an undirected network, represented by a symmetric adjacency matrix, it may be questionable to reach a directed network (with an asymmetric adjacency matrix) as the outcome of the planner's optimal intervention. One desirable property that we might ask is the preservation of symmetry.

## 4.2 Case 2: Symmetric intervention

In this case  $\mathcal{A} = \{M \in \mathcal{M}_{n,n} \text{ such that } M = M^T\}$ . The following results is another direct application of Higham, 2000 ([?]):

**Proposition 7.** The closest matrix - in the Frobenius norm - in  $\mathcal{A} = \{M \in \mathcal{M}_{n,n} : M = M^T\}$  to  $A$  such that the action profile  $\bar{\mathbf{a}}$  is played in the game played on the transformed network is:

$$A + E_{\min}^{Sym}(\bar{\mathbf{a}}) \quad (4.6)$$

with

$$E_{\min}^{Sym}(\bar{\mathbf{a}}) = \frac{(b - A\bar{\mathbf{a}})\bar{\mathbf{a}}^T + \bar{\mathbf{a}}^T(b - A\bar{\mathbf{a}})^T}{\bar{\mathbf{a}}^T \bar{\mathbf{a}}} - (b - A\bar{\mathbf{a}})^T \bar{\mathbf{a}} \frac{\bar{\mathbf{a}} \bar{\mathbf{a}}^T}{\bar{\mathbf{a}}^T \bar{\mathbf{a}}} \quad (4.7)$$

**Remark 2.** <sup>37</sup> -The intervention of the planner is a rank-2 matrix -It is possible to show that the cost of intervention  $\mu^{Sym}$  is close to the cost of the intervention without constraint

$$\mu \leq \mu^{Sym} \leq \sqrt{2}\mu$$

---

<sup>37</sup>See Higham (2000)

### 4.3 Case 3: Sparse intervention

In some situations the planner might not be able to modify some features of the network. For example she might not be able to create a link between player  $\tilde{i}$  and  $\tilde{j}$ . One important case is when we do not allow for self-loops, that is  $g_{ii} = 0$ , for any  $i$ .

Let  $Y$  be a matrix such that  $y_{ij} \in \{0, 1\}$ ,  $\forall i, j$ . We define the following set:

**Definition 3.**

$$S_2(Y) = \{M \in \mathcal{M}_{n,n} \text{ such that } M_{ij} = 0 \text{ if } Y_{ij} = 0\}$$

From Higham, 2000 ([?]) we know that, given  $Y$  and  $S_2$  the solution to the minimization problem (4.2) with  $\mathcal{A} = S_2(Y)$  is:

$$E_{\min}^{Sp}(\bar{\mathbf{a}}) = \sum_{i=1..n} \frac{1}{(\bar{\mathbf{a}}_i^T \bar{\mathbf{a}})} \epsilon_i^T (b - A\bar{\mathbf{a}}) \epsilon_i s_i^T$$

where  $\bar{\mathbf{a}}_i \in \mathbb{R}^n$ , with  $j$ -th element  $\bar{a}_i(j)$  defined as:

$$\bar{a}_i(j) = \begin{cases} \bar{\mathbf{a}}(j) & \text{if } Y_{ij} = 1 \\ 0 & \text{if } Y_{ij} = 0 \end{cases}$$

For example we want to study the case where the initial and post-intervention do not have self-loops (that is  $a_{kk} = 0$  for all  $k$ ), we define the restrictions matrix:

$$B = \begin{bmatrix} 0 & 1 & 1 & \dots \\ 1 & 0 & 1 & \dots \\ & \dots & & \\ \dots & 1 & 1 & 0 \end{bmatrix}$$

Consequently we get:

$$E_{\min}^{Sp}(\bar{\mathbf{a}}) = \begin{bmatrix} \frac{b_1 - \epsilon_1 A \bar{\mathbf{a}}}{\sum_{i \neq 1} a_i^2} & 0 & 0 & \dots \\ 0 & \frac{b_2 - \epsilon_2 A \bar{\mathbf{a}}}{\sum_{i \neq 2} a_i^2} & 0 & \dots \\ & \dots & & \\ \dots & 0 & 0 & \frac{b_n - \epsilon_n A \bar{\mathbf{a}}}{\sum_{i \neq n} a_i^2} \end{bmatrix} \begin{bmatrix} 0 & a_{12} & a_{13} \dots \\ a_{21} & 0 & a_{23} & \dots \\ & \dots & & \\ \dots & a_{n,n-2} & a_{n,n-1} & 0 \end{bmatrix}$$

**Remark 3.** This intervention is a rank- $n$  intervention. As noted in Dennis and Schnabel, 1979 ([?]) this

*type of correction is of no computational significance and can be made "one row at the time". Column  $k$  of the intervention is proportional to the non-zero entries to the objective equilibrium profile  $k$ -th component.*

Dennis and Schnabel, 1979 ([?]) and other work in the computer science study other interesting case of  $\mathcal{A}$ . In the next section we focus on studying a problem that we consider of particular interest: How should a planner, with a given budget, intervene on the network structure in order to increase total welfare.

## 5 Closest network structure that maximizes welfare

In the previous section we specified how a social planner can implement an equilibrium profile at minimal cost. While this might be of interest per se, in this section, we will use the results we obtained to address an important policy question: how should the social planner allocate resources to maximize total welfare.

### 5.1 The planner's problem

The goal of the social planner is to maximize aggregate welfare knowing that agents are utility-maximizers and that they will play the Nash equilibrium  $\mathbf{a}^* \in \mathbb{R}^n$  of the game described in the model. To reach this goal the social planner can intervene on the network structure provided she respects a cost constraint. Given  $A \in \mathcal{M}_{nn}$  (as defined in (3.1) as a function of the initial network structure  $G$ ), a budget  $C > 0$ , and a vector of characteristics  $b \in (\mathbb{R}_+^*)^n$ , the planner's problem  $\mathcal{P}[A, C, b]$  is the following:

$$\begin{aligned} \max_{E \in \mathcal{M}_{nn}} \quad & \sum_{i \in \mathcal{N}} W_i [\mathbf{a}^*(E)] \\ \text{s.t.} \quad & [A + E] \mathbf{a}^*(E) = b, \\ & \|E\|_F^2 \leq C \end{aligned} \quad (\mathcal{P}[A, C, b])$$

where  $\mathbf{a}^*(E)$  is the Nash equilibrium of the game played on the transformed matrix  $A + E$ .

### 5.2 Equilibrium profile's reaction to planner's intervention

Even if the solution to problem  $(\mathcal{P}[A, C, b])$  is very sensible to the initial conditions, we try to describe what is the general idea behind the planner's intervention. In particular we will try to compare the equilibrium profile of the network game before and after intervention. In order to this we will project the equilibrium

on a particular system of coordinates and observe how these projections are affected before and after the intervention of the planner. We will show how an appropriate choice of the coordinates system can help us understand which player will have the incentives more affected by the planner's intervention and why.

We will first rewrite the problem of the planner in a equivalent form. After that we will recall some notions of matrix algebra that will use in Theorem 1 to get our result.

We use the result of proposition (6) to express the problem of the planner in  $(\mathcal{P}[A, C, b])$ :

Given  $b \in \mathbb{R}^n$ ,  $A \in \mathcal{M}_{n,n}$ ,  $C > 0$ ,

$$\begin{aligned} \max_{\mathbf{a} \in \mathbb{R}^n} \quad & \frac{\|\mathbf{a}\|^2}{2} \\ \text{s.t.} \quad & C\|\mathbf{a}\|^2 - \|A\mathbf{a} - b\|^2 \geq 0 \end{aligned} \quad (\mathcal{P}_2[A, C, b])$$

We used a classical result of the quadratic cost network game literature, for which at the Nash equilibrium  $\mathbf{a}^*$ :<sup>38</sup>

$$W = \sum_{i \in \mathcal{N}} W_i(\mathbf{a}^*) = \frac{\|\mathbf{a}^*\|^2}{2} \quad (5.1)$$

A solution to  $(\mathcal{P}_2[A, C, b])$  exists as the objective function is continuous and the constrained set is compact. Applying the extreme value theorem yields existence of a solution. We call  $\tilde{\mathbf{a}}^*$  a solution of  $(\mathcal{P}_2[A, C, b])$ .

The idea is to compare the equilibrium of the initial game, with the equilibrium of the game after the intervention of the planner. The equilibria we want to compare are n-dimensional vectors and therefore a metric for comparison is difficult to obtain. What we will do is to choose an appropriate set of coordinates and try to compare the projections of the two equilibria on these. In order to do this we will have to recall some notion of matrix algebra.

**Singular value decomposition** In order to analyze the changes that the equilibrium action profile is going through when the planner's intervention takes place, we introduce a common tool of linear algebra,

---

<sup>38</sup>see Bramoullé, Kranton and D'Amours, 2014, [?]

the singular value decomposition. Given a matrix  $M \in \mathcal{M}_{n,n}$  there exists a factorization, called singular value decomposition (SVD) of  $M$  of the form

$$M = U\Sigma V^T,$$

where  $U$  and  $V$  are unitary matrices of  $\mathcal{M}_{n,n}$  and  $\Sigma$  is a diagonal matrix with non-negative real numbers on the diagonal. The diagonal entries  $\{s_i\}_i$  of  $\Sigma$  are known as the singular values of  $M$ , the column of  $U$  (or  $V$ ) are known as left (or right) singular vectors of  $M$ .

**Case of symmetric positive definite matrices** If  $M$  is symmetric and positive definite, its singular value decomposition coincides with its eigendecomposition. In this case  $U = V$ , the column of the matrix  $U$  are the eigenvectors of  $M$  and the singular values are its associated eigenvalues. The advantage of the singular value decomposition over the eigendecomposition is that it always exists.

Let us consider the singular value decomposition of  $A$ :

$$A = U\Sigma V^T \tag{5.2}$$

with  $u^i$  (respectively  $v^i$ ) the  $i$ -th column of  $U$  (respectively  $V$ ), and  $s_i$  the  $i$ -th singular value, when ranking the singular values in decreasing order.  $u_j^i$  (respectively  $v_j^i$ ) is the  $j$ -th element of the vector  $u^i$  (respectively  $v^i$ ). Let  $p_i$  be the projection of the initial equilibrium action profile  $\mathbf{a}^*$  (before intervention) on  $v^i$  and  $\tilde{p}_i$  the projection of the new action profile  $\tilde{\mathbf{a}}^*$  (after intervention).

We add an assumption about the size of the budget. This condition is sufficient for our next result to hold but not necessary.

**Assumption 4.**

$$C < \frac{\left(\min_i s_i^2\right)^2}{\sum_{i=1}^n s_i^2}$$

The following theorem allows us to rank the ratio of any two projections before and after intervention.

**Theorem 1.** *Under assumption 4, for  $i, j = 1, \dots, n$ ,  $i < j$ :*

$$\frac{p_i}{p_j} \leq \frac{\tilde{p}_i}{\tilde{p}_j} \quad (5.3)$$

*Proof.* See section 9.1 of appendix. For the version where symmetry is imposed to the transformed network, see section 9.3 of appendix.

The ratio of the projections of the resulting action profile on different right singular vectors of  $A$  tells us how close the action profile is from one right singular vector relative to another one. By comparing this ratio before and after intervention, we understand whether the action profile moves towards one right singular vector relative to another one with the intervention. The next corollary formalizes this idea.

For any  $\mathbf{a} \in \mathbb{R}^n$ , let  $\theta_i(\mathbf{a})$  be the angle between the vector  $\mathbf{a}$  and  $v^i$  in the 2-dimensional subspace of  $\mathbb{R}^n$  spanned by  $\mathbf{a}$  and  $v^i$ .

**Corollary 2.** *Under assumption 4, for  $i, j = 1, \dots, n$ ,  $i < j$ :*

$$\frac{\theta_i(\mathbf{a}^*)}{\theta_j(\mathbf{a}^*)} \leq \frac{\theta_i(\tilde{\mathbf{a}}^*)}{\theta_j(\tilde{\mathbf{a}}^*)} \quad (5.4)$$

*Proof.* See appendix.

Corollary (2) means that the action profile moves towards  $v^i$  and away from  $v^j$ , for all  $i < j$ .

This result holds in the particular case when the singular value decomposition is the eigendecomposition as well as in the general case when it is not. But the interpretation of the result is easier when the  $\{v^i\}_i$  are the  $n$  eigenvectors of  $A$ . In this case it exists the eigenvectors of  $A$  and of  $G$  are the same and because the eigenvectors of an adjacency matrix have a nice interpretation in terms of network structure, we can give an interpretation to our result.

Remember the relationship between  $A$  and  $G$  from (3.1):

$$A = I - \beta G$$

As  $I$  is a diagonal matrix, the correspondence between eigenvectors and eigenvalues of the two matrices is straightforward and depends on the sign of  $\beta$ . Let's call  $\{\lambda_i\}_i$  the eigenvalues of  $A$  and  $\{\lambda_i(G)\}_i$  the eigenvalues of  $G$ .

**Lemma 3.**

$$\lambda_i(G) = \frac{1 - \lambda_i}{\beta} \quad \text{with } v^i \text{ the associated eigenvector} \quad (5.5)$$

and the  $n$  eigenvectors  $\{v^i\}_{i=1}^n$  of  $G$  are the eigenvectors of  $A$  ranked:

1. If  $\beta > 0$ : in the opposite order as  $\{v^i\}_{i=1}^n$
2. If  $\beta < 0$ : in the same order as  $\{v^i\}_{i=1}^n$

*Proof.* From (3.1) we get the following equivalence:

$$Av^i = \lambda_i v^i \quad \Leftrightarrow \quad Gv^i = \frac{1 - \lambda_i}{\beta} v^i \quad (5.6)$$

The results directly follows.

The order of the eigenvectors refers to the orders of their associated eigenvalues ranked in decreasing order, from the largest to smallest. The eigenvectors of the adjacency matrix capture important characteristic of the network. We will use these to interpret the result on the projections of the equilibria. Each component of the first eigenvector of the adjacency matrix represents the eigenvector centrality of the corresponding player. A high eigenvector centrality means that a node is connected to many nodes who themselves have high centrality. The higher the  $j$ -element of the first eigenvector is, the more central the  $j$ -th player is. The last eigenvector, instead, in a bipartite network, assign negative values to players in one of the two sets and positive values to the one in the other set.<sup>39</sup> We put together corollary 2 with lemma 3 to give a result that has a clear interpretation in terms of network structure:

**Proposition 8.** *Under assumption 4, for  $i, j = 1, \dots, n$ ,  $i < j$ :*

1. If  $\beta > 0$ :

*The equilibrium responds to the planner's intervention moving **from the higher-ranked to the lower-ranked eigenvector** in the subspace of  $\mathbb{R}^n$  spanned by those two eigenvectors.*

---

<sup>39</sup>A bipartite network is a network whose nodes can be divided into two disjoint and independent sets  $U$  and  $V$  such that every edge connects a vertex in  $U$  to one in  $V$ .



2. If  $\beta < 0$ :

*The equilibrium responds to the planner's intervention moving **from the lower-ranked to the higher-ranked eigenvector** in the subspace of  $\mathbb{R}^n$  spanned by those two eigenvectors.*

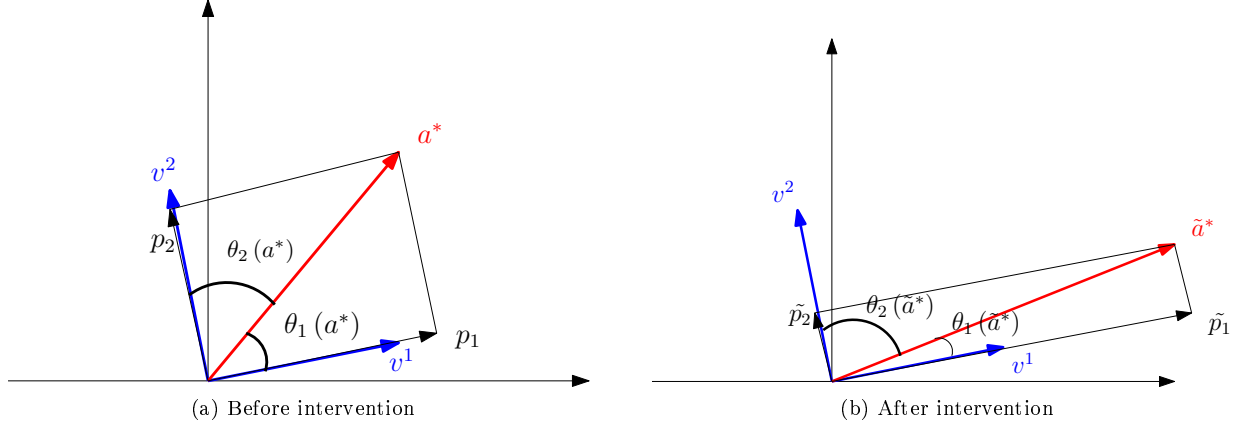


Figure 1: Change in the relative distance between the action profile and the first two right singular vectors of  $A$

Figure (1) is a graphical representation of corollary (2), for  $i = 1, j = 2$  in the case of strategic complements. Proposition (8) says that more central agents (of the initial network), in this case player 1, as his component of the first eigenvector is larger, increase more their action relative to another weighting of the agents (the weighting described by  $\bar{v}^2$  for instance). This result tells us that the planner is changing the incentives in the game in such a way that central players are the one more affected. They will respond more than others to the intervention. We see the decision of the planner to target central player as a result of the importance of these players typical of network games of strategic complements. When a central player, after intervention, increases his action he incentivizes the players that are connected to him to increase their action as well, bringing benefits to all the populations. The more a player is central, the more his action is important to incentivize other players. In section 7 we will try to investigate how the social planner actually targets central players, but first we want to explain what happens in game of strategic substitutes.

When we consider a game of strategic substitutes the two graphs of Figure 1 follow the opposite order. To get an interpretation we focus on the last eigenvector of the network. Proposition 8 says that the equilibrium will tend to mirror the last eigenvector (with respect to any other eigenvector) after the intervention of the planner. To fix ideas consider Figure (2). Here all the nodes are connected but some links are stronger (dark blue) and form a bipartite graph. The last eigenvector of the adjacency matrix will have positive entries for

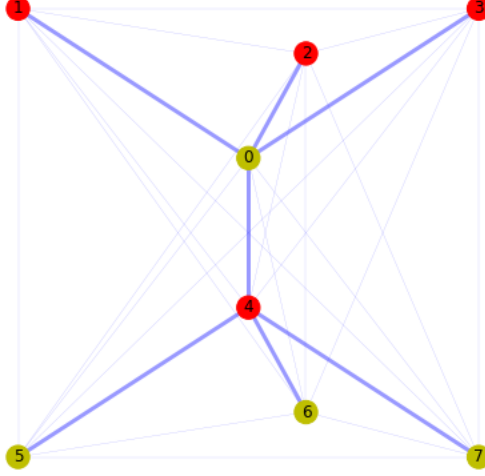


Figure 2: Bipartite Network

the red nodes and negative for the green. Proposition 8 is telling us that the social planner is changing the incentives of red and green players in opposite directions. Red players, after interventions will increase their actions while green will decrease them. Changing actions of closed neighbors in opposite direction increases total welfare. If the planner would instead move incentives in the same direction an increase in the action of a player would crowd out the incentive of his neighbor. In section 7 we try to understand how the planner reaches her goal. Before, in the next section, we will compare our result to GGG.

## 6 Comparison with GGG

We now want to compare our result with the proposition 1 of GGG. We focus on the special case where the singular value decomposition is the eigendecomposition. The results follow through in terms of singular vectors. We find that the changes in the action profile have the same **direction** in terms of eigenvectors, though the **amplitude** of the changes is surely different.

**Definition 4.**  $q^i(G, b)$  is the projection of  $b$  on the  $i$ -th eigenvector of  $G$ .

The result of proposition 1 of GGG is:

1. If  $\beta > 0$ ,

$$\frac{q^l(G, b^{\text{new}}) - q^l(G, b)}{q^l(G, b)} \text{ is weakly decreasing in } l \quad (6.1)$$

2. If  $\beta < 0$ ,

$$\frac{q^l(G, b^{\text{new}}) - q^l(G, b)}{q^l(G, b)} \text{ is weakly increasing in } l \quad (6.2)$$

where  $b$  is the initial vector of individual characteristics, and  $b^{\text{new}}$  is the new vector, after intervention (remember that they intervene on  $b$  when we intervene on  $G$ ).

In order to compare their result and our result, let us rename our projections:

**Definition 5.**  $p^i(b, G, C)$  is the projection of the action profile solution of  $(\mathcal{P}_2[A, C, b])$ , with parameters  $b \in \mathbb{R}^n$ ,  $G \in \mathcal{M}_{n,n}$ ,  $C > 0$ , on the  $i$ -th eigenvector of  $G$

Using simple algebra and Proposition 1 of GGG we can state:

**Theorem 2.** For  $i < j$ :

(1) (6.1) implies

$$\frac{p^i(b, G, C)}{p^j(b, G, C)} \leq \frac{p^i(b^{\text{new}}, G, C)}{p^j(b^{\text{new}}, G, C)} \quad (6.3)$$

(2) (6.2) implies

$$\frac{p^i(b, G, C)}{p^j(b, G, C)} \geq \frac{p^i(b^{\text{new}}, G, C)}{p^j(b^{\text{new}}, G, C)} \quad (6.4)$$

In words, it means that the results they find when  $\beta > 0$  and  $\beta < 0$  leads to the same direction of the change in the action profile  $\mathbf{a}^*$  with respect to any two eigenvectors of  $G$  as in our result of proposition 8. Note that the inequalities (6.3) and (6.4) are inequalities, telling us nothing about the amplitude of the variation. But the variation in  $\mathbf{a}^*$  (our object of interest as it determines aggregate welfare) goes in the same direction in both cases of  $\beta$ .

The two different policies used by the planner move the incentives of the players in a similar fashion. Proposition 1 in GGG describes how the planner achieves her goal moving the private benefits of the players. In the next section we describe how the planner achieves her goal changing the structure of the network.

## 7 Network structure analysis

In section 6 we saw how the social planner change the incentives of the players, studying how the equilibrium moves after the interventions. In this section we want to give insights on how she actually achieves her goal, in terms of which connections will be modified and how.

Following the notations of equilibrium profiles, we call  $\tilde{A} = A + E_{\min}$  (how the matrix  $A$  is transformed after intervention) and  $\tilde{G}$  the new network structure. As we will study the object  $E_{\min}$ , we rename it  $\Delta A$  as it represents what is added to the matrix  $A$  in the optimal intervention. Similarly we define  $\Delta G = \tilde{G} - G$ . From the definition (3.1) we get:

$$\tilde{G} - G = -\frac{1}{\beta}(\tilde{A} - A) \quad (7.1)$$

Therefore studying  $\tilde{A} - A = \Delta A$  sheds light on the change of network structure  $\Delta G$ . We focus on undirected network so that we can use the eigendecomposition and have an easy way to interpret the results. Furthermore we do not set any restriction on the planner intervention. We will study separately how the planner intervene on:

1. links arriving to a set of players
2. links starting from a set of players

Following the approach of the previous sections, we work on the projection of  $\Delta A$  on the basis  $B$  of the eigendecomposition of  $A$ :

$$B = \{(v_1 v_1^T), (v_1 v_2^T), \dots, (v_1 v_n^T), (v_2 v_1^T), \dots, (v_n v_1^T), \dots, (v_n v_n^T)\} \quad (7.2)$$

The decomposition is a sum of  $n^2$  elements:

$$\Delta A = \sum_{i=1, \dots, n, j=1, \dots, n} \mu_{i,j} \left[ v^i (v^j)^T \right] \quad (7.3)$$

We define the following two objects of interest:

**Definition 6.** For  $i = 1, \dots, n$ :

$$\begin{aligned} S^I(v^i) &= \sum_{k=1}^n (v^i)^T \Delta A v^k \\ S^O(v^i) &= \sum_{k=1}^n (v^k)^T \Delta A v^i \end{aligned} \tag{7.4}$$

$S^I(v^i)$  is the sum of the coefficients of decomposition (7.3) on the basis elements of  $B$  with vector  $v^i$  as left vector of the outer product.  $S^O(v^i)$  is the equivalent measure when  $v^i$  is the right vector of the outer product. With this quantity we want to capture how relevant is the eigenvector  $v^i$  to explain the structure of  $\Delta A$  (the in-links are captured with  $S^I(v^i)$  and the out-links with  $S^O(v^i)$ ).

**Proposition 9.** When  $C \rightarrow 0$ , for any  $i, j$ :

$$\frac{S^I(v^i)}{S^I(v^j)} = \left( \frac{1 - \beta \lambda_j(G)}{1 - \beta \lambda_i(G)} \right)^2 \frac{v_i^T b}{v_j^T b} \tag{7.5}$$

and

$$\frac{S^O(v^i)}{S^O(v^j)} = \left( \frac{1 - \beta \lambda_j(G)}{1 - \beta \lambda_i(G)} \right) \frac{v_i^T b}{v_j^T b} \tag{7.6}$$

with  $\lambda_i(G)$  defined in equation (5.5) as a function of the eigenvalues of  $A$  in the following way:

$$\lambda_i(G) = \frac{1 - \lambda_i}{\beta}$$

where  $\lambda_i(G)$  is:

1. the  $(n - i + 1)$ -th largest eigenvalue of  $G$  if  $\beta > 0$
2. the  $i$ -th largest eigenvalue of  $G$  if  $\beta < 0$

*Proof.* See appendix

Proposition 9 gives us information about  $\Delta A$  but we are eventually interested in  $\Delta G$ . Recalling (7.1), we notice that  $\Delta G$  and  $\Delta A$  are proportional, and thus :

$$\frac{\sum_{k=1}^n (v^i)^T \Delta G v^k}{\sum_{k=1}^n (v^j)^T \Delta G v^k} = \frac{S^I(v^i)}{S^I(v^j)} \quad (7.7)$$

As a consequence we directly use the ratio  $\frac{S^I(v^i)}{S^I(v^j)}$  to inform us about  $\Delta G$ . To avoid ambiguity, let us call  $r_i$  the  $i$ -th largest eigenvalue of  $G$  regardless of the sign of  $\beta$ , and  $w_i$  its associated eigenvector. Therefore, when  $\beta > 0$ ,  $r_1 = \lambda_n(G)$  and  $r_p = \lambda_{n-p+1}(G)$ . Those notations are merely relabeling to provide clearer intuition of proposition 9, as when  $\beta > 0$ ,  $\lambda_i(G)$  is the  $n - i + 1$ -th largest eigenvalue of  $G$ , but  $v^i$  stays its associated eigenvector, that is  $v^i = w^{n-i+1}$ . We can then rewrite the expressions of proposition 9 in a more intuitive way in the following corollary.

**Corollary 3.** *When  $C \rightarrow 0$ , for  $p > q$ :*

$$\frac{S^I(w^q)}{S^I(w^p)} = \left( \frac{1 - \beta r_p}{1 - \beta r_q} \right)^2 \frac{w_q^T b}{w_p^T b} \quad (7.8)$$

and

$$\frac{S^O(w^q)}{S^O(w^p)} = \left( \frac{1 - \beta r_p}{1 - \beta r_q} \right) \frac{w_q^T b}{w_p^T b} \quad (7.9)$$

**Interpretation for strategic complements ( $\beta > 0$ )** In the case of strategic complements we would like to explain how the social planner incentivize central players, the result we obtained from section 6. We call

$$W(p) = \frac{w_1^T b}{w_p^T b} \quad (7.10)$$

and  $R(p)$ :

$$R(p) = \left( \frac{1 - \beta r_p}{1 - \beta r_1} \right) \quad (7.11)$$

We can thus rewrite the results of corollary 7.8, when  $q = 1$  as:

$$\frac{S^I(w^1)}{S^I(w^p)} = R(p)^2 W(p) \quad \text{and} \quad \frac{S^O(w^1)}{S^O(w^p)} = R(p) W(p) \quad (7.12)$$

The larger is the first of the ratio in (7.12) the more the planner is increasing the intensity of links toward

central players. The larger the second of the two ratio in (7.12) the more the planner is increasing the intensity of links from central players to the other individuals. We can now study the two ratio and make some interesting observations:

- From assumption 2 we know that both the numerator and the denominator of  $R(p)$  are between 0 and 1 (and as  $r_1 \geq r_p$ ), as  $|r_1 - r_p|$  grows the more the social planner is targeting links from and towards central players. Other thing beeing equal we can compare networks that have different eigenvalues and understand how much centrality of players has a key role in the decision of the planner.
- As  $R(p) \geq 1$  (strict inequality for  $r_1 \neq r_p$ ),  $R(p)^2$  grows faster than  $R(p)$  with  $|r_1 - r_p|$ , and thus the ratio for in-links grows faster than the ratios for out-links.
- The inner product  $\langle w^i, b \rangle$  describes how high the private benefits of the action is for the players that matter in the  $i$ -th direction of the network structure. For instance,  $w^1$  singles out central players and  $\langle w^1, b \rangle$  is high when central players have high private marginal benefits from the action. The more central players have high private value for the action relative to players important according the  $p$ -th dimension, the more  $W(p)$  is large. The idea is that the planner has an incentive to induce links with players that have high private benefits for the action as a link with one of this player is more valuable than a link with a similar player with lower private benefit. In other words  $W(p)$  gives us the relative size of externalities in the direction of the structural aspect described by eigenvector  $w^1$  versus by eigenvector  $w^p$ .
- $W(p)$  can either be smaller or bigger than 1.
  - If  $W(p) > 1$ , then  $S^k(w^1)/S^k(w^p)$ ,  $k = I, O$  are both strictly greater than 1, (this ) which means that the planner focuses more budget on central players relative to the players singled out by the dimension  $w^p$ . The effects contained in the ratios  $R(p)$  and  $W(p)$  reinforce each other.
  - On the other hand, if  $W(p) < 1$ , the effects of  $R(p)$  and  $W(p)$  go compensate each other. The first ratio pushes the planner to focus on central players, but on the other hand those players do not have a high marginal return on the action (indicated by  $W(p) < 1$ ) and the planner may want to split her budget between central players and players with high private returns.

**Interpretation for strategic substitutes ( $\beta < 0$ )** We showed in section 6 that the social planner, in the case of strategic substitutes, moves incentive of neighbors' players in opposite directions. We try to understand how she achieves this goal.

We denote by  $R^*(q)$ :

$$R^*(q) = \left( \frac{1 - \beta r_n}{1 - \beta r_q} \right) \quad (7.13)$$

We can thus rewrite the results of corollary 7.8, when  $p = n$  as:

$$\frac{S^I(w^q)}{S^I(w^n)} = R^*(q)^2 \frac{w_q^T b}{w_n^T b} \quad \text{and} \quad \frac{S^O(w^q)}{S^O(w^n)} = R^*(q) \frac{w_q^T b}{w_n^T b} \quad (7.14)$$

As opposed to the complement case,  $R^*(q) \leq 1$  and decreases as the distance between  $r_q$  and  $r_n$  increases. The smaller is  $R^*(q)$  the more the planner lower the intensity of links between green players and links between red players in Figure (2). In this way the network will tend to become a bipartite network. Green players will invest a lot in the action and red players will do the opposite.

## 8 Other Interventions

Started from the results given in Section 4, it is possible to investigate other interesting problem for the planner, as, for example, how to modify the network in order to give a minimum utility to all the players or how to guarantee some form of equality among players or again how to increase the utility (or action) of all the players by the same amount. The social planners, in the case where  $a$  represents effort might also be interested in maximizing total effort. We leave these and other questions for future studies.



## 9 Appendix

### 9.1 Proof of theorem 1

We remind that the mathematical definition of the projection of any vector  $x \in \mathbb{R}^n$  on the vector  $v^i$  is:

$$\text{projection of } x \text{ on } v^i = \langle x, v^i \rangle \quad (9.1)$$

The steps of the proof are the following: first we derive an expression for  $p_i$  by using the singular value decomposition of  $A$ . Then, we derive properties on the equilibrium profile after optimal intervention  $\tilde{\mathbf{a}}^*$  by differentiating the Lagrangian of the optimization problem. From there we get an expression for  $\tilde{p}_i$ . Finally we use the expressions for  $p_i$  and  $\tilde{p}_i$  to deliver a ranking of the ratios of projections at different indices.

#### 9.1.1 Projection of the equilibrium profile before intervention

The equilibrium before intervention is a vector  $\mathbf{a}^*$  such that

$$A\mathbf{a}^* = b \quad (9.2)$$

Pre-multiplying both sides of the equation by the vector  $(v^i)^T A^T$  we get:

$$(v^i)^T A^T A\mathbf{a}^* = (v^i)^T A^T b \quad (9.3)$$

**Right hand side of (9.3)** We are now using the singular value decomposition of  $A$ :  $U\Sigma V^T$ . As  $(v^1, \dots, v^n)$  forms an orthonormal basis of  $\mathbb{R}^n$ , we have the following equality:

$$(v^i)^T V = \begin{bmatrix} \langle v^i, v^1 \rangle & \langle v^i, v^2 \rangle & \dots & \langle v^i, v^n \rangle \end{bmatrix} = (e_i^n)^T \quad (9.4)$$

where  $e_i^n$  is the  $i$ -th vector of the Euclidean basis of  $\mathbb{R}^n$ , that is its  $i$ -th element is 1 and all its other elements are 0. This helps us to simplify  $(v^i)^T A^T$ :

$$(v^i)^T A^T = (v^i)^T V \Sigma U^T = s_i (u^i)^T \quad (9.5)$$

By post-multiplying by the vector of characteristics  $b$  we reach:

$$(v^i)^T A^T b = s_i (u^i)^T b \quad (9.6)$$

**Left hand side of (9.3)**

$$(v^i)^T A^T A = s_i (u^i)^T A \quad (9.7)$$

(using the equality in (9.5)). Re-using the singular value decomposition, we get:

$$(v^i)^T A^T A = s_i (u^i)^T U \Sigma V^T \quad (9.8)$$

Similarly to (9.4) we have

$$(u^i)^T U = \begin{bmatrix} \langle u^i, u^1 \rangle & \langle u^i, u^2 \rangle & \dots & \langle u^i, u^n \rangle \end{bmatrix} = (e_i^n)^T \quad (9.9)$$

This helps us to simplify (9.8):

$$(v^i)^T A^T A = s_i (e_i^n)^T \Sigma V^T = s_i^2 (v^i)^T \quad (9.10)$$

By post-multiplying by  $\mathbf{a}^*$  we reach:

$$(v^i)^T A^T A \mathbf{a}^* = s_i^2 (v^i)^T \mathbf{a}^* \quad (9.11)$$

**Expression for the projection of  $\mathbf{a}^*$  on  $v^i$**  Putting together left-hand side and right-hand side of (9.3), we reach an expression for  $(v^i)^T \mathbf{a}^*$ , that is  $p_i$ :

$$p_i = \langle v^i, \mathbf{a}^* \rangle = (v^i)^T \mathbf{a}^* = \frac{(u^i)^T b}{s_i} \quad (9.12)$$

**Note on the symmetric case (undirected network)** When the initial matrix  $G$  (and therefore  $A$ ) is symmetric:

1.  $v_i = u_i$  is an eigenvector of  $G$  (and therefore of  $A$ )
2.  $s_i = \lambda_i$  is the eigenvalue  $i$ -th of the matrix  $A$

Hence:

$$v_i^T a^* = \frac{v_i^T b}{\lambda_i} \quad (9.13)$$

### 9.1.2 Projection of the equilibrium profile after intervention

Let us first derive conditions on the solution  $\tilde{a}^*$  of  $(\mathcal{P}_2[A, C, b])$  by writing and differentiating the lagrangian  $\mathcal{L}(\cdot)$  of the problem.

$$\begin{aligned} \mathcal{L}(a, \mu) &= \frac{\|a\|^2}{2} + \mu \left[ C\|a\|^2 - \|Aa - b\|^2 \right] \\ \frac{\partial \mathcal{L}}{\partial a}(a, \mu) &= a + \mu 2ca - \mu (2A^T Aa - 2A^T b) \end{aligned} \quad (9.14)$$

Setting this expression equal to zero we get:

$$(1 + 2\mu c)\tilde{a}^* - 2\mu A^T A\tilde{a}^* = -2\mu A^T b \quad (9.15)$$

Dividing both sides by  $-2\mu$  (as the constraint is binding and thus the multiplier is strictly positive) and calling  $K = \frac{1}{2\mu} + c$  we get :

$$-K\tilde{a}^* + A^T A\tilde{a}^* = A^T b \quad (9.16)$$

Pre-multiplying both sides by  $(v^i)^T$  and directly using the results of (9.6) and (9.10), we reach

$$-Kv_i^T \tilde{a}^* + s_i^2 (v^i)^T \tilde{a}^* = s_i (u^i)^T b \quad (9.17)$$

and finally we have an expression for the projection of  $\tilde{a}^*$  on the vector  $v^i$ :

$$\tilde{p}_i = v_i^T \tilde{a}^* = \frac{s_i}{s_i^2 - K} (u^i)^T b \quad (9.18)$$

### 9.1.3 Comparing the ratio of projections on two different right singular vectors

From (9.13) and (9.18) we compute the ratio of interest:

$$\frac{\tilde{p}_i/\tilde{p}_j}{p_i/p_j} = \frac{s_i^2 s_j^2 - K}{s_j^2 s_i^2 - K} = \frac{s_i^2}{s_i^2 - K} \bigg/ \frac{s_j^2}{s_j^2 - K} \quad (9.19)$$

Let's observe that assumption (4) together with the Kuhn-Tucker conditions of (  $\mathcal{P}_2[A, C, b]$  ) tells us that  $s_i^2 - K > 0 \quad \forall i$ . In order to establish this consider the FOCs of the lagrangian:

$$\tilde{\mathbf{a}}^* + \mu 2C \tilde{\mathbf{a}}^* - \mu 2A^T A \tilde{\mathbf{a}}^* = -\mu 2A^T b$$

Simple algebra gives:

$$\left( \frac{1}{2\mu} + C \right)^2 = \frac{(A\tilde{\mathbf{a}}^* - b)^T A A^T (A\tilde{\mathbf{a}}^* - b)}{\|\tilde{\mathbf{a}}^*\|^2} = \frac{\|A^T (A\tilde{\mathbf{a}}^* - b)\|^2}{\|\tilde{\mathbf{a}}^*\|^2} \quad (9.20)$$

Cauchy-Schwartz inequality tells us that for a matrix  $M \in \mathcal{M}_{n,n}$ , a vector  $x \in \mathbb{R}^n$ ,  $\|Mx\|_2 \leq \|M\|_F \|x\|_2$ . We apply this result (note also that the Frobenius norm of a matrix and that of its transposed are the same) to our previous equality:

$$(9.20) \Rightarrow \left( \frac{1}{2\mu} + C \right)^2 \leq \frac{\|A\tilde{\mathbf{a}}^* - b\|^2 \|A\|^2}{\|\tilde{\mathbf{a}}^*\|^2} \quad (9.21)$$

Besides, we know that the constraint of (  $\mathcal{P}_2[A, C, b]$  ) is binding:

$$\frac{\|A\tilde{\mathbf{a}}^* - b\|^2}{\|\tilde{\mathbf{a}}^*\|^2} = C$$

Therefore, plugging this into (9.21), we get:

$$\left( \frac{1}{2\mu} + C \right)^2 \leq C \|A\|_F^2$$

Which is equivalent to:

$$K^2 \leq C \|A\|_F^2 = C \sum_{i=1}^n s_i^2 \quad (9.22)$$

As  $K$  is  $> 0$  (the lagrangian multiplier is  $> 0$ ), we get the following:

$$K \leq \sqrt{C} \sqrt{\sum_{i=1}^n s_i^2} \leq \frac{\min_i s_i^2}{\sqrt{\sum_{i=1}^n s_i^2}} \sqrt{\sum_{i=1}^n s_i^2} = \min_i s_i^2 \quad (9.23)$$

where the second inequality comes from assumption (4)

(9.23) implies, for all  $i$ :

$$s_i^2 \geq K \quad (9.24)$$

We now use this result to determine whether the ratio in (9.19) is bigger or smaller than 1. The function  $f : x \mapsto \frac{x}{x-K}$  is increasing in  $x$  iff  $x > K$  (which we know from (9.24)). This implies that

$$(9.20) \Rightarrow \frac{\tilde{p}_i/\tilde{p}_j}{p_i/p_j} > 1 \quad \text{iff} \quad s_i > s_j \quad (9.25)$$

In the symmetric case the projection on the left singular vectors corresponds to the projection on the eigenvector of the adjacency matrix.

#### 9.1.4 Proof of Corollary (2)

Corollary (2) comes directly from the fact that the projection of  $\mathbf{a}^*$  on  $v^i$  can also be written as:

$$p_i = \langle \mathbf{a}^*, v^i \rangle = \|\mathbf{a}^*\| \|v^i\| \cos \theta_i(\mathbf{a}^*) \quad (9.26)$$

Besides, the vectors of the family  $\{v^i\}_i$  are orthonormal, therefore the norm of each of them is 1. The ratio of projections thus boils down to:

$$\frac{p_i}{p_j} = \frac{\cos \theta_i(\mathbf{a}^*)}{\cos \theta_j(\mathbf{a}^*)} \quad \text{and} \quad \frac{\tilde{p}_i}{\tilde{p}_j} = \frac{\cos \theta_i(\tilde{\mathbf{a}}^*)}{\cos \theta_j(\tilde{\mathbf{a}}^*)} \quad (9.27)$$

We pause a moment to consider the sign of the projections. In principle,  $\langle \mathbf{a}^*, v^i \rangle$  can be either negative or positive (depending whether the angle  $\theta_i(\mathbf{a}^*)$  is bigger or smaller than  $90^\circ$ ). We can abstract from this sign ambiguity by "choosing" the (right) singular vector that makes the projection positive. If a vector  $x$  is a singular vector of a matrix with norm 1, then the vector  $-x$  will also be a singular vector of this matrix (provided we change the sign of both the left and the right singular vectors associated to the same singular value), of norm 1 too. Therefore we choose all the right singular vectors of  $A$  such that  $\langle \mathbf{a}^*, v^i \rangle > 0$  for all  $i$ . Then, by continuity of the inner product, for a budget small enough,  $\langle \tilde{\mathbf{a}}^*, v^i \rangle$  will also be positive.

Then, combining (9.27) and theorem 1 yields the desired result.

## 9.2 Proof of proposition 9

**Step 1: expression for  $\tilde{\mathbf{a}}^*$**  From (9.16), we get:

$$\tilde{\mathbf{a}}^* = (A^T A - KI)^{-1} A^T b \quad (9.28)$$

(remember that  $K$  is a scalar, equal to  $\frac{1}{2\mu} - C$  and that the matrix  $A^T A - KI$  is invertible as it is symmetric) . Going back to the eigendecomposition of  $A$ , and taking into account that  $A$  is symmetric and thus that  $A = A^T$ , we have:

$$(A^T A - KI) = A^2 - KI = V \Sigma^2 V^T - KI = V \begin{pmatrix} \lambda_1^2 - K & & \\ & \ddots & \\ & & \lambda_n^2 - K \end{pmatrix} V^T \quad (9.29)$$

It is then easy to compute  $(A^T A - KI)^{-1}$ :

$$(A^T A - KI)^{-1} = V \begin{pmatrix} \frac{1}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & \frac{1}{\lambda_n^2 - K} \end{pmatrix} V^T \quad (9.30)$$

Finally:

$$\tilde{\mathbf{a}}^* = (A^T A - KI)^{-1} V \Sigma V^T b = V \begin{pmatrix} \frac{\lambda_1}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & \frac{\lambda_n}{\lambda_n^2 - K} \end{pmatrix} V^T b \quad (9.31)$$

**Step 2: expression for  $(\mathbf{b} - A\tilde{\mathbf{a}}^*)$**  Another useful factorization is the one of  $(\mathbf{b} - A\tilde{\mathbf{a}}^*)$ . Directly from the previous result in (9.31), we get:

$$(\mathbf{b} - A\tilde{\mathbf{a}}^*) = \mathbf{b} - AV \begin{pmatrix} \frac{\lambda_1}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & \frac{\lambda_n}{\lambda_n^2 - K} \end{pmatrix} V^T b = \left( I - V \begin{pmatrix} \frac{\lambda_1^2}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & \frac{\lambda_n^2}{\lambda_n^2 - K} \end{pmatrix} V^T \right) b$$

or again

$$(\mathbf{b} - A\tilde{\mathbf{a}}^*) = V \begin{pmatrix} 1 - \frac{\lambda_1^2}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & 1 - \frac{\lambda_n^2}{\lambda_n^2 - K} \end{pmatrix} V^T \mathbf{b} \quad (9.32)$$

**Step 3: expression for  $(v^i)^T \Delta A v^j$**  We now compute  $(v^i)^T \Delta A v^j$  for all  $i, j$  by using the two previous steps together with (4.4) of proposition 6 (with  $\bar{\mathbf{a}} = \mathbf{a}^*$ ):

$$(v^i)^T \Delta A v^j = (v^i)^T \frac{(\mathbf{b} - A\tilde{\mathbf{a}}^*)(\tilde{\mathbf{a}}^*)^T}{(\tilde{\mathbf{a}}^*)^T \tilde{\mathbf{a}}^*} v^j \quad (9.33)$$

whose right-hand side is equal to:

$$\frac{1}{(\tilde{\mathbf{a}}^*)^T \tilde{\mathbf{a}}^*} (v^i)^T V \begin{pmatrix} 1 - \frac{\lambda_1^2}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & 1 - \frac{\lambda_n^2}{\lambda_n^2 - K} \end{pmatrix} V^T \mathbf{b} \mathbf{b}^T V \begin{pmatrix} \frac{\lambda_1}{\lambda_1^2 - K} & & \\ & \ddots & \\ & & \frac{\lambda_n}{\lambda_n^2 - K} \end{pmatrix} V^T v^j$$

or again is equal to

$$\frac{1}{(\tilde{\mathbf{a}}^*)^T \tilde{\mathbf{a}}^*} \frac{-K}{\lambda_i^2 - K} (v^i)^T \mathbf{b} \mathbf{b}^T \frac{\lambda_j}{\lambda_j^2 - K} v^j$$

finally we get:

$$(v^i)^T \Delta A v^j = -\frac{1}{(\tilde{\mathbf{a}}^*)^T \tilde{\mathbf{a}}^*} \frac{K \lambda_j}{(\lambda_i^2 - K)(\lambda_j^2 - K)} \langle v^i, \mathbf{b} \rangle \langle v^j, \mathbf{b} \rangle \quad (9.34)$$

**Step 4: computing  $S^I(v^i)$**  Now we want to compute the sum  $S^I(v^i)$  from previous result:

$$S^I(v^i) = \sum_{j=1}^n (v^i)^T \Delta A v^j = -\frac{1}{(\tilde{\mathbf{a}}^*)^T \tilde{\mathbf{a}}^*} \langle v^i, \mathbf{b} \rangle \frac{1}{\lambda_i^2 - K} \sum_{j=1}^n \frac{K \lambda_j}{(\lambda_j^2 - K)} \langle v^j, \mathbf{b} \rangle \quad (9.35)$$

**Step 5: computing the ratio  $S^I(v^1)/S^I(v^p)$**  For  $p = 2, \dots, n$ :

$$\frac{S^I(v^1)}{S^I(v^p)} = \frac{\langle v^1, \mathbf{b} \rangle \frac{\lambda_p^2 - K}{\lambda_1^2 - K}}{\langle v^p, \mathbf{b} \rangle \frac{\lambda_1^2 - K}{\lambda_p^2 - K}} \quad (9.36)$$

**Step 6: taking the limit as  $C \rightarrow 0$**  Following the previous approach, we take  $C \rightarrow 0$  and work for a transformed network close to the initial network. We know from (9.22) that:

$$K^2 \leq C \|A\|_F^2 = C \sum_{i=1}^n \lambda_i^2$$

Therefore when  $C \rightarrow 0$ ,  $K \rightarrow 0$  too and we get:

$$\frac{S^I(v^1)}{S^I(v^p)} = \frac{\langle v^1, b \rangle \lambda_p^2}{\langle v^p, b \rangle \lambda_1^2} \quad (9.37)$$

**Step 7: Ratio as a function of the eigenvalues of  $G$**  The final step is again to transform expression (9.41) as a function of the eigenvalues of  $G$  and not of  $A$ . Using the correspondence stated in lemma 3, we get:

$$\frac{S^I(v^1)}{S^I(v^p)} = \frac{\langle v^1, b \rangle}{\langle v^p, b \rangle} \left( \frac{1 - \beta \lambda_p(G)}{1 - \beta \lambda_1(G)} \right)^2 \quad (9.38)$$

However it is to be noted that when  $\beta > 0$ , the order of the eigenvalues of  $A$  and  $G$  is inverted and therefore  $\lambda_1(G)$  is the **smallest** eigenvalue of  $G$  and  $\lambda_n(G)$  is the **largest** eigenvalue of  $G$

**Step 8: computation of  $S^O(v^i)$**  Remember that

$$S^O(v^i) = \sum_{k=1}^n (v^k)^T \Delta A v^i$$

We restart from (9.34) above where we just change the index over which we sum:

$$S^O(v^j) = -\frac{1}{(\tilde{\mathbf{a}}^*)^T \tilde{\mathbf{a}}^*} \frac{\lambda_j}{\lambda_j^2 - K} \langle v^j, b \rangle \sum_{i=1}^n \frac{K}{(\lambda_i^2 - K)} \langle v^i, b \rangle \quad (9.39)$$

**Step 9: computing the ratio  $S^O(v^1)/S^O(v^p)$**  For  $p = 2, \dots, n$ :

$$\frac{S^O(v^1)}{S^O(v^p)} = \frac{\langle v^1, b \rangle \lambda_p^2 - K \lambda_1}{\langle v^p, b \rangle \lambda_1^2 - K \lambda_p} \quad (9.40)$$

Following the previous approach, we take  $C \rightarrow 0$  and can thus neglect the constant  $K$  and we get:

When  $C \rightarrow 0$ :

$$\frac{S^O(v^1)}{S^O(v^p)} = \frac{\langle v^1, b \rangle \lambda_p}{\langle v^p, b \rangle \lambda_1} \quad (9.41)$$



**Step 10: Ratio as a function of the eigenvalues of  $G$**  The final step is again to transform expression (9.41) as a function of the eigenvalues of  $G$  and not of  $A$ . Using the correspondence stated in lemma 3, we get:

$$\frac{S^O(v^1)}{S^O(v^p)} = \frac{\langle v^1, b \rangle}{\langle v^p, b \rangle} \frac{1 - \beta \lambda_p(G)}{1 - \beta \lambda_1(G)} \quad (9.42)$$

However it is to be noted that when  $\beta > 0$ , the order of the eigenvalues of  $A$  and  $G$  is inverted and therefore  $\lambda_1(G)$  is the **smallest** eigenvalue of  $G$  and  $\lambda_n(G)$  is the **largest** eigenvalue of  $G$ .

This proves the result.

### 9.3 Optimal intervention under symmetry constraint

From (4.7) we have that

$$E_{\min}(\bar{a}) = \frac{(b - A\bar{a})\bar{a}^T + \bar{a}^T(b - A\bar{a})}{\bar{a}^T \bar{a}} - (b - A\bar{a})^T \bar{a} \frac{\bar{a}\bar{a}^T}{\bar{a}^T \bar{a}}$$

Let us first derive conditions on the solution  $\tilde{a}^*$  of the planner's problem under symmetry constraint by writing and differentiating the Lagrangian  $\mathcal{L}(\cdot)$  of the problem.

$$\begin{aligned} \mathcal{L}(a, \mu) &= (a^T a) + \mu \left[ C(a^T a)^2 - 2(b - Aa)^T (b - Aa) a^T a + ((b - Aa)^T a)^2 \right] \\ \frac{\partial \mathcal{L}}{\partial a}(a, \mu) &= a + 2\mu c(a^T a)a - 2\mu(a^T a)(2A^2 a - 2Ab) - 2\mu(w^T w)a + 2\mu(w^T a)(b - Ay) \end{aligned} \quad (9.43)$$

where we pose

$$w \equiv b - Ay$$

Rearranging we get

$$[(1 + 2\mu c(a^T a) - 2\mu(w^T w))I - 4\mu(a^T a)A^2 - 2\mu(w^T a)A]a = [-4\mu(a^T a)A - 2\mu(w^T a)I]b$$

This give us an expression for the new equilibrium profile action (that we call  $a^*$  in this section only). Using the properties of eigenvectors-eigenvalues, as in the case examined before we get:

$$a^* = V \begin{bmatrix} \frac{-4\mu(a^T a)\lambda_1 - 2\mu(w^T a)}{(1 + 2\mu c(a^T a) - 2\mu(w^T w)) - 4\mu(a^T a)\lambda_1^2 - 2\mu(w^T a)\lambda_1} & \dots & \frac{-4\mu(a^T a)\lambda_n - 2\mu(w^T a)}{(1 + 2\mu c(a^T a) - 2\mu(w^T w)) - 4\mu(a^T a)\lambda_n^2 - 2\mu(w^T a)\lambda_n} \end{bmatrix} V^T b \quad (9.44)$$

where  $\lambda_i$  is the  $i$ -th eigenvalue of  $A$

Dividing numerator and denominator by  $2\mu(a^T a)$ , calling  $\gamma = -\frac{w^T a}{a^T a} > 0$  and  $K = \frac{1}{2\mu a^T a} + c - \frac{w^T w}{a^T a}$  we get:

$$a^* = V \begin{bmatrix} \frac{2\lambda_1 - \gamma}{2\lambda_1^2 - \gamma\lambda_1 - K} & \dots & \frac{2\lambda_n - \gamma}{2\lambda_n^2 - \gamma\lambda_n - K} \end{bmatrix} V^T b \quad (9.45)$$

Projecting on eigenvector  $v^i$ :

$$a^* v^i = \frac{2\lambda_i - \gamma}{2\lambda_i^2 - \gamma\lambda_i - K} (v^i)^T b$$

As before I want to compare the ratio of projection on different eigenvectors before and after intervention. This boils down to the study of the function

$$f(\lambda_i) = \frac{2\lambda_i^2 - \gamma\lambda_i}{2\lambda_i^2 - \gamma\lambda_i - K}$$

When  $C \rightarrow 0$ ,  $\gamma \rightarrow 0$  as  $w \rightarrow 0$  and  $K > 0$ . This yields the result. For  $K > 0$ , multiply by  $a^T$  the equation of the derivative of the lagrangian and divide by  $2\mu a^T a$ . This gives us

$$K = a^T (2A^2 - \gamma A) (a - A^{-1}b)$$

which is  $>0$  when  $\gamma$  is small enough

# Chapter 3

## Stopping contagion: optimal network intervention

### 1 Introduction

We study contagion processes on social networks. We investigate how a social planner should optimally intervene on the network structure to prevent them. Many welfare-relevant phenomena can be described as contagion processes in networks. The most studied one is epidemics, in this case we investigate which kind of prevention programs the planner should promote. Other interesting applications are the diffusion of bad rumors and fake news, or risky behaviors such as crime and smoking. We ask in all those cases which preventive measures the planner could take. <sup>40</sup>

We use the Susceptible-Infected-Susceptible (SIS) model from the epidemiology literature as a convenient way to address such processes.<sup>41</sup> Individuals are in one of two possible states: Susceptible S or Infected I. The probability that a susceptible individual becomes infected is increasing in the number of individuals he interacts with who are infected. The probability that an infected individual becomes susceptible again is exogenous and given by a parameter. This gives tractability and fits well the epidemics example. <sup>42</sup> The key characteristics of those diffusion processes is that my behavior or my state (sick or sound) may evolve over time, and the transition from one state to the other may depend on the states of the individuals I interact with. For instance, I am more likely to get infected by a disease if I meet individuals that are infected themselves, and the more such individuals I meet, the more likely it becomes. Recently the SIS model has been used to understand a large number of processes. In information diffusion the biggest the number of my friends knowing about a rumor, the more probable it is that I learn it, and afterwards transmit it. In coordination games, my best response is to cooperate if my friends do so and to cheat if they do.

In this environment we work on long run outcomes and study the steady-state of the system, with the interpretation that this represents the fraction of time each individual spends in the infected state over a

---

<sup>40</sup>The empirical literature on peer effects shows that teenagers are more inclined to start smoking if their friends do (Robalino and Macy, 2018, [?]).

<sup>41</sup>see Pastor-Satorrás and Vespignani, 2001, [?])

<sup>42</sup>It may become a limitation for other contagion processes. In the context of social conventions, the probability to switch from one convention to another seems to depend on the convention my friends adopted in any of the two directions. We leave the study of cases where the transition between one state to the other is symmetric across the states for future work.

long period of time.<sup>43</sup>

In this classical SIS framework we ask how a social planner could intervene to prevent an outbreak. We allow the planner to decrease the probability that an infection is transmitted when a meeting with an infected individual occurs. The planner could be a politician that promotes a campaign to increase the use of protective measures to decrease the spread of sexually transmitted diseases. She could notify social media users about dubious sources of information in order to prevent the spread of fake news. All these intervention are costly, and the larger is the intervention the higher is the cost the planner has to sustain. Of course multiple interventions constitute possible way to stop the epidemic diffusion, so we ask how to reach our goal at minimal cost. How does the policy of the social planner depends on the structure of interactions among players? How to distribute resources in the population?

The first important contribution of the paper is the characterization of the intervention of the planner (Corollary 5). We managed to have an analytical solution to the problem that does not require any limitation on the initial structure of the network. This result gives the social planner a useful guidance when deciding how to allocate resources. It is important to stress that network-nature problem like the one we face, where we have to deal with a large amount of information, are very difficult to solve and require a lot of computational power. Therefore providing a closed-form solution becomes of even larger importance.

To prove our result we borrow from the literature on epidemiology and from the one of computer science. In fact, we show that, mathematically, the problem corresponds to the transformation of the matrix representing the interactions across individuals into its "closest" negative semidefinite matrix.<sup>44</sup> This problem is well studied in the computer science literature, we could therefore apply some of the techniques commonly used in that field. While the main result is a direct application of one of these, we are the first to use them to study diffusion processes in networks.

Even if the interventions would differ a lot depending on the initial structure of the network, we argue that the planner intervenes in a systematic way. She tries to decrease interactions across different communities

---

<sup>43</sup>In these models the outcomes of interest are divided between short run and long run. We focus on the second one. It is to be noted that the steady-state has a different meaning when working on approximate versus exact networks. In the first case, the system is deterministic and once the steady-state is reached, the system does not move anymore. The steady-state then describes the fraction of individuals of each type (for instance their degree) in the infected state at each period. Of course some individuals change state at each periods but for each type, individuals leaving the infected state are replaced by the same quantity leaving the susceptible state, provided the population is large enough. In the case of exact networks, each individual is unique and thus the system is stochastic and its state (which is a  $n$ -dimensional vector collecting the states of each individuals for a population of size  $n$ ) changes at each period. The steady-state can be seen as a measure of time spent in the Infected state.

<sup>44</sup>See Higham (2000) for a survey of the literature.

in the population and she focus on eliminating the disease in each community separately.

If, on one side, the result is unique in the literature for its generality, on the other side, there is an important limitation. The optimal intervention might result in making an interaction with an infected individual decrease the probability to spread contagion. This means that the planner's prevention program should not only make the probability of contagion smaller than the initial one but it should be able to promote immunization from interaction across individuals. While this would make sense if the infected state has the flavor of a rival good, the interpretation is not straightforward when we think about diseases. Aware of this limitation, in example 1 and example 2, we show that when we analyze two of the most-studied networks' structures in the literature, we should not be concerned with the problem. Furthermore, we argue that, even when the optimal intervention requires immunization to increase from some interactions, our result is useful. In fact, Theorem 3 gives us the exact cost of the planner's optimal program. If we only allow her to decrease the probability of contagion from the initial one, our measure of the cost gives a useful lower bound on the total expenditure the planner needs to prevent the spread of the disease.

In section 5 we study two families of networks. First, we consider a situation where players differ in their degrees. Some players have an higher number of connections (degree) with respect to others. We compare populations with different degree distributions. We want to compare a population where all the players are similar (have the same degree) with an heterogeneous population (some players have an high degree and other a small one). Using Corollary 5 we show that it is easier for the planner to prevent the spread of the disease in the first case. In proposition 10 we extend this result showing that, for comparable networks, the easiest scenario to face for the social planner is one where all the players are connected among them. Finally, we compare populations with two communities that are more or less integrated between them. We show that the social planner's optimal strategy is to limit interactions between individuals of the two communities no matter what is the initial configuration. We argue that these interactions are crucial to the spread of the disease in the population and that is why are the target of the planner. Fighting the disease separately in the two communities is the best way for the planner to defeat the disease.

## 2 Literature Review

Our paper inscribes itself in the literature on diffusion processes among a population of individuals who interact with one another.<sup>45</sup> There are two main modeling choices that impact the properties of those types of processes: the structure of connections between individuals and the details of how the transmission takes place from one individual to another. The SIS model has been examined under a wide range of connection structures. In the benchmark version, the pattern of interactions is not fixed and individuals have the same probability of meeting any other individual at each period. This is called the homogeneous mixing assumption. A refined version of this model is one where individuals are defined by the number of connections they have, their degree. The homogeneous mixing assumption is maintained but some individuals meet more people than other. Consequently the outcome varies by type of individuals. Lopez-Pintado, 2007 ([?]) and Jackson and Rogers, 2007 ([?]) are examples of this approach. The pattern of connections is thus expressed through the degree distribution of the population, that is, the fraction of individuals having exactly  $d$  friends for all possible  $d$ . This modeling choice yields tractability but fails to grasp various important features of connection patterns, such as geography or the long-term aspect of interactions. This is why we chose to study exact arbitrary networks. In exact networks, every individual is different in principle and has his own set of links.

Regarding the transmission function, Gleeson, 2013, ([?]) compares diffusion processes under different contagion mechanisms. The probability of contagion can be an increasing function of the absolute number of my friends who are infected (SIS model), of the relative number of my friends who are infected (voter model), it can follow a majority rule, a threshold rule, or an Ising Glauber model.<sup>46</sup> Beyond those well-defined transmission function, Lopez-Pintado, 2007 ([?]) brings an interesting contribution to the literature by working with an unspecified function. Lopez-Pintado, 2007 varies the properties of this function and analyzes how her outcomes of interest change. As opposed to those papers, we chose the SIS model for its tractability as it allowed us to deal with more complexity on the network structure side, which is our focus.

While it is really interesting to analyze the dynamic of the SIS model, we decide to focus on long run outcomes (Steady-States). We do this for tractability and to compare our work to the economic literature. Among the papers studying the steady-state of contagion processes, we distinguish two main objects of interest. A strand of literature (like Jackson and Rogers, 2007, [?]) focuses on positive steady-states (where

---

<sup>45</sup>To see a full review on modeling dynamical systems on networks, see Porter and Gleeson, 2016 ([?])

<sup>46</sup>See Gleeson (2013) for a discussion

at least some individuals or some types have a strictly positive steady-state value) and derive properties of those steady-states. For instance, Jackson and Rogers 2007 provides comparative statics on the average level of infection as a function of the dispersion of the degree distribution representing the network. Another branch of literature explores when the steady-state is null (nobody is infected, an initial seeding of a disease dies out before spreading) versus strictly positive (an outbreak of the epidemics occurs). Lopez-Pintado, 2007 ([?]) is an example.

We belong to this last group, but as opposed to most of this literature that focuses on which characteristics of the process (ratio of the individual transmission and remission rate) allow to prevent an outbreak for different network structures, we take the parameters of the process as given and we ask which network structure reaches the zero-steady-state for those parameters. Galeotti and Rogers (2013) consider an intervention where part of the individuals get vaccination and therefore cannot be infected. Differently from them the planner in our case targets links and individuals. While Galeotti and Rogers (2013) considers a specific type of network we try to give a result that generalize to all networks' structures.

### 3 The Model

In this section we present the SIS model and we recall some important results on epidemic diffusion from the literature. The aim of this section is to understand which kind of network configurations favor the diffusion of epidemics and which one do not in the SIS framework. While the model was originally developed with the aim of explaining epidemiological applications it is now used to study application such as information diffusion, learning and imitation dynamics.

We study a population of  $n$  individuals located on a network. The set of player is called  $\mathcal{N} = 1, \dots, n$ . The network is described by the adjacency matrix  $G \in \mathcal{M}_{n,n}(\mathcal{R}^+)$ . Each element  $g_{ij}$  of  $G$  represents the intensity of the link between individual  $i$  and  $j$ .<sup>47</sup> <sup>48</sup>

We model the epidemic process in continuous time. At each time  $t$ , each individual can either be susceptible, or infected. Let  $X_i(t)$  be the Bernoulli random variable that takes value 1 if node  $i$  is infected at time  $t$  and value 0 if it is not. An infected node may become susceptible at a constant rate  $\delta > 0$ . The infection rate of a susceptible node  $i$  is  $\lambda \sum_{j=1}^n g_{ij} X_j(t)$ . Here  $\lambda$  is a parameter measuring the contagiousness of the disease. The second term of the product  $\sum_{j=1}^n g_{ij} X_j(t)$  capture the importance of the interaction with the other players. The probability of becoming infected is increasing in the probability of nodes  $j$  being infected weighted by the link that  $i$  shares with  $j$ .

Following Pastor-Satorras et al (2015) [?], we write the equations governing the evolution of the expectation of  $X_i(t)$  (which is also the probability that node  $i$  is infected at time  $t$  as  $X_i(t)$  is Bernoulli).

$$\frac{dE[X_i(t)]}{dt} = E \left[ -\delta X_i(t) + (1 - X_i(t)) \lambda \sum_{j=1}^n g_{ij} X_j(t) \right] \quad (3.1)$$

The term inside the expectation on the RHS of equation (3.1) when  $i$  is infected is equal to  $-\delta$  (the recovery rate) while it is equal to  $\lambda \sum_{j=1}^n g_{ij} X_j(t)$  (the probability of infection) when  $i$  is susceptible.

Since  $X_i(t)$  is a Bernoulli we can rewrite equation (3.1) as:

---

<sup>47</sup>In the literature there are two different approaches to model networks. It is possible to impose that all connections among players have the same intensity. In this case the adjacency matrix has only entrances 0 or 1.  $g_{ij}$  is 1 when there is a link between  $i$  and  $j$ , 0 otherwise. We decide, instead, to model the network using a weighted adjacency matrix. Connections between players can vary in intensity. We believe that this approach fits well when analyzing the diffusion of diseases. This approach is also necessary for the results we obtain

<sup>48</sup>We limit the study to symmetric networks. Links are bidirectional.  $i$  can be affected by  $j$  with the same probability as  $j$  can be infected by  $i$ .



$$\frac{dx_i(t)}{dt} = -\delta x_i(t) + \lambda \sum_{j=1}^n g_{ij} x_j(t) - \lambda \sum_{j=1}^n g_{ij} E[X_i(t)X_j(t)] \quad (3.2)$$

where  $x_i(t)$  the probability that  $i$  is infected at time  $t$ .

### 3.1 Discussion of the epidemic threshold

It is a well known result that equation (3.2) has always one steady-state where  $x_i(t) = 0$  for all  $i$ . Sometimes a positive steady-state exists as well. A crucial goal of the epidemics literature is to determine a threshold for the diffusion parameters  $\lambda, \delta$  such that the initial seed of the disease does not result in an outbreak, or again such that the zero-steady-state is the only one. We will recall two important results from the literature on epidemics (Lemma 4 and Lemma 5) that explain the role of the largest eigenvalue of the adjacency matrix  $G$  on the existence of an outbreak (non-zero steady state).

#### Epidemic threshold lower bound in the exact SIS model

**Lemma 4.** *Let  $\lambda_1$  be the largest eigenvalue of  $G$ . A lower bound for the threshold of the epidemic process  $\frac{\lambda}{\delta}$  in the exact version of the SIS model is  $\frac{1}{\lambda_1}$ :*

$$\frac{\lambda}{\delta} \leq \frac{1}{\lambda_1} \quad \Rightarrow \quad \text{there is no outbreak} \quad (3.3)$$

*Proof.* Following Pastor-Satorras et al (2015) [?], we revisit equation (3.2) and note that for all  $i$ ,  $\sum_{j=1}^n g_{ij} X_i(t)X_j(t) \geq 0$ . We can thus transform (3.2) into an inequality by removing the last term, and replacing expectations of Bernoulli random variables by their probability of success:

$$\frac{dx_i(t)}{dt} \leq -\delta x_i(t) + \lambda \sum_{j=1}^n g_{ij} x_j(t) \quad (3.4)$$

This inequality holds for all  $i$ . We create a system of  $n$  inequalities,  $i = 1, \dots, n$ . We observe that setting inequalities to equations, the system boils down to (3.12), whose solution has been found in (3.19). We deduce that:

$$\mathbf{x}(t) \leq \sum_{r=1}^n a_r(0) e^{(\lambda \lambda_r - \delta)t} \mathbf{v}_r \quad (3.5)$$

The fastest growing term (as  $t$  increases) of (3.5) is the one associated with the highest positive eigenvalue  $\lambda_1$ . This expression shows that in order to get  $\mathbf{x}(t)$  going to 0, we need all exponential factors to be negative, or again:

$$\frac{\lambda}{\delta} \leq \frac{1}{\lambda_1} \quad \Rightarrow \quad \text{the right hand side of (3.5) decays exponentially} \quad (3.6)$$

Therefore the inverse of the highest eigenvalue of the adjacency matrix  $G$  is a lower bound for the threshold of the epidemic process (meaning  $\frac{\lambda}{\delta}$ ) at which the disease does not degenerate in an outbreak in the exact model.

□

This first result tells us that a sufficient condition not to have an outbreak is that the largest eigenvalue of the adjacency matrix has to be small enough with respect to the ratio  $\frac{\delta}{\lambda}$ . Intuitively this ratio measure how fast an individual recovers from the disease ( $\delta$ ) with respect to how contagious the disease is ( $\lambda$ ). We would like the implication in (3.6) to be a double implication. While this is not always true, we show how adding and additional assumption gives us the result.

**SIS-epidemic threshold under individual-based mean-field approximation (IBMF)** We assume that the two random variables of neighboring nodes are uncorrelated. While it seems a strong assumption, this is the standard in the literature.<sup>49</sup> We will keep this assumption throughout the paper.

**Assumption 5.**

$$E[X_j(t)X_i(t)] = E[X_j(t)]E[X_i(t)] \quad \text{for all } t, i, j \quad (3.7)$$

Under assumption 5 it is possible to prove the following

**Lemma 5.** *The disease dies out and an epidemic is avoided in the IBMF approximation of the SIS model iff:*

$$\frac{\lambda}{\delta} \leq \frac{1}{\lambda_1} \quad (3.8)$$

---

<sup>49</sup>Pastor-Satorrar et al. (2015) study the accuracy of the individual mean-field approximation for different type of networks.

*Proof.* We remind the equations governing the evolution of the expectation of  $X_i(t)$ :

$$\frac{dE[X_i(t)]}{dt} = E \left[ -\delta X_i(t) + (1 - X_i(t)) \lambda \sum_{j=1}^n g_{ij} X_j(t) \right] \quad (3.9)$$

We use an individual-based mean-field approximation (IBMF), assuming that the status of neighboring nodes are independent, or again:

Using (3.7) in (3.9), and replacing the expectation of the Bernoulli variables by their probability, we get:

$$\frac{dx_i(t)}{dt} = -\delta x_i(t) + (1 - x_i(t)) \lambda \sum_{j=1}^n g_{ij} x_j(t) \quad (3.10)$$

We choose the initial conditions so that at  $t = 0$  we have a small number  $c$  of infected individuals and everyone else is susceptible, so that  $x_i(0) = c/n$ . One way to get insight from (3.10) is to see that as  $n$  grows large,  $1 - x_i(0)$  goes to 1, and following Pastor-Satorras et al (2015) [?], we approximate (3.10) by replacing  $1 - x_i(0)$  by its limit value:

$$\frac{dx_i(t)}{dt} = -\delta x_i(t) + \lambda \sum_{j=1}^n g_{ij} x_j(t) \quad (3.11)$$

or in matrix form:

$$\frac{d\mathbf{x}(t)}{dt} = \lambda A(G) \mathbf{x}(t) \quad (3.12)$$

with  $\mathbf{x}(t)$  the vector of the  $\{x_i(t)\}_i$  and:

$$A(G) = G - \frac{\delta}{\lambda} I \quad (3.13)$$

To find a solution to the system of differential equations in (3.12), we decompose  $\mathbf{x}(t)$  on the orthonormal basis composed of the eigenvectors of  $A$  (which are the same as the eigenvectors of  $G$ ) that we call  $\{v_r\}_{r=1}^n$  where  $v_r$  is the eigenvector associated with  $\lambda_r$ , the  $r$ -th eigenvalue of  $G$ , where eigenvalues are ranked in decreasing order. We call  $a_r(t)$  the coefficient of the decomposition associated to  $v_r$ :

$$\mathbf{x}(t) = \sum_{r=1}^n a_r(t) v_r \quad (3.14)$$

Differentiating this expression we get:

$$\frac{d\mathbf{x}(t)}{dt} = \sum_{r=1}^n \frac{a_r(t)}{dt} v_r$$

Combining the previous equation with (3.12):

$$\sum_{r=1}^n \frac{da_r(t)}{dt} v_r = \lambda A \mathbf{x}(t) = \lambda A \sum_{r=1}^n a_r(t) v_r = \lambda \sum_{r=1}^n a_r(t) A v_r \quad (3.15)$$

and finally, using the fact that  $\lambda_r$  is the eigenvalue of  $G$  associated with the eigenvector  $v_r$  and that from (3.13) we can write the  $r$ -th eigenvalue of  $A$  as  $\lambda_r - \frac{\delta}{\lambda}$ :

$$\sum_{r=1}^n \frac{da_r(t)}{dt} v_r = \lambda \sum_{r=1}^n a_r(t) \left( \lambda_r - \frac{\delta}{\lambda} \right) v_r \quad (3.16)$$

As  $\{v_r\}_{r=1}^n$  constitutes a basis of  $\mathbb{R}^n$ , the decomposition of any vector on it is unique and thus,  $\forall r, t$ :

$$\frac{da_r(t)}{dt} = a_r(t) (\lambda \lambda_r - \delta) \quad (3.17)$$

For each  $r, t$ , the above differential equation has for solution:

$$a_r(t) = a_r(0) e^{(\lambda \lambda_r - \delta)t} \quad (3.18)$$

Finally, plugging (3.19) into the decomposition (3.14), we get:

$$\mathbf{x}(t) = \sum_{r=1}^n a_r(0) e^{(\lambda \lambda_r - \delta)t} v_r \quad (3.19)$$

The same argument as in the previous proof tells us that:

$$\lambda_1 \leq \frac{\delta}{\lambda} \quad (3.20)$$

□

## 4 Optimal immunization

We determined in the previous section the relationship between the largest eigenvalue of  $G$  and the epidemic threshold under different approximations. The inverse of the highest eigenvalue of  $G$ ,  $\frac{1}{\lambda_1}$  is the exact epidemic

threshold in the IBMF approximation, and the upper bound of this threshold if we remove assumption 5.

In this section we model the intervention of the planner and we state the main result of the paper.

We assume that the social planner can intervene to change the structure of the network determining the diffusion of the disease. If  $g_{i,j}$  is the weight that determine the probability that  $i$  is infected from  $j$  the planner can invest resources to lower this probability to  $g'_{i,j} < g_{i,j}$ . We make the following assumption about the cost she incurs to alter the network:

**Assumption 6.** *Changing the structure of the network from  $G$  to  $G^*$  has a cost of  $\|G^* - G\|_F$ , where  $\|M\|_F = \sqrt{\sum_{ij} m_{ij}^2}$  is the Frobenius norm of  $M$ .*

The previous assumption captures the idea that the social planner can modify the interaction between  $i$  and  $j$  at a convex cost. The planner faces increasing marginal costs of intervention.<sup>50</sup>

Once we established a metric on the cost of intervention of the planner we can ask what is the least costly intervention that prevent the diffusion of the disease in the population. As we discussed in the previous section this is equivalent to impose conditions so that in steady state all the variable measuring infection are equal to zero. Lemma 5 gives us a good starting point to solve the problem of the planner. Having the highest eigenvalue of  $G$  less or equal to  $\frac{\delta}{\lambda}$  is equivalent to having the matrix  $A(G)$  (defined in (3.13)) be semi-definite negative. As a result we have that:

**Corollary 4.** *The least costly intervention for the planner is the one that change the network from  $G$  to  $G'$ , where  $G'$  is the solution to:*

$$\min\{\|A(G) - A(G')\|, \text{ such that } A(G') \text{ is negative semi-definite}\}$$

This corollary enables us to resort to a famous result of the computer science literature: we use the theorem 2.1 of Higham(1998)[?] to find the nearest symmetric negative semi-definite matrix of  $A(G)$ , and to see how to perform this manipulation. Higham(1998) [?] also gives us a closed form solution for the cost of intervening on any given network structure, as a function of  $\frac{\lambda}{\delta}$ . While the result comes directly using the technique adopted by Higham(1998) [?] the application to the study of epidemic diffusion is completely new.

---

<sup>50</sup>The Frobenius norm can be seen as an extension of the euclidean norm to  $\mathcal{R}^{n \times n}$

In order to state the theorem, we remind that  $A(G)$  and  $G$  have the same eigenvectors  $\{v_r\}_{r=1}^n$ . We call  $\{\mu_r\}_{r=1}^n$  the  $n$  eigenvalues of  $A(G)$ ,  $\mu_r$  being associated with the eigenvector  $v_r$ . They can be expressed as a function of the eigenvalues of  $G$  in the following way:

$$\mu_r = \lambda_r - \frac{\delta}{\lambda} \quad (4.1)$$

The following theorem specify the optimal intervention of the planner:

**Theorem 3.** *The planner intervenes on the network changing  $A(G)$  to  $\tilde{A}(G)$ , where:*

$$\tilde{A}(G) = V\Delta V^T \quad (4.2)$$

with  $V$  the matrix whose  $r$ -th column is the  $r$ -th eigenvector of  $G$ ,  $v_r$ , and where  $\Delta$  is the following diagonal matrix:

$$\Delta = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

with

$$d_r = \begin{cases} \mu_r & \text{if } \lambda_r < \frac{\delta}{\lambda} \\ 0 & \text{otherwise} \end{cases}$$

The cost to reach it from  $A$  is:

$$\|A - \tilde{A}\|_F^2 = \sum_{\lambda_r > \frac{\delta}{\lambda}} \left( \lambda_r - \frac{\delta}{\lambda} \right)^2 \quad (4.3)$$

*Proof.* The result directly follows from theorem 2.1 of Higham(1998)[?] and its proof.  $\square$

We can restate the theorem in term of the new network  $G$  instead of  $A(G)$  to have a better interpretation of the result

**Corollary 5.** *The closest network structure  $\tilde{G}(G)$  of  $G$  such that no outbreak occurs is:*

$$\tilde{G}(G) = V\Delta V^T + \frac{\delta}{\lambda}I = V\Lambda V^T \quad (4.4)$$

with

$$\Lambda = \begin{bmatrix} \tilde{\lambda}_1 & & \\ & \ddots & \\ & & \tilde{\lambda}_n \end{bmatrix}$$

and

$$\tilde{\lambda}_r = \begin{cases} \lambda_r & \text{if } \lambda_r < \frac{\delta}{\lambda} \\ \frac{\delta}{\lambda} & \text{otherwise} \end{cases}$$

We are changing the all the eigenvalues of the matrix  $G$  that exceed the ratio  $\frac{\delta}{\lambda}$ , while leaving unchanged the ones below this ratio. All the eigenvectors of the original adjacency matrix remained unchanged. The first comment on the property of the new network structure regards what is known in the network literature as the spectral gap:

**Definition 7.** *The spectral gap is the difference between the largest and the second largest eigenvalues of the adjacency matrix.*

A straightforward application of 5 is:

**Corollary 6.** *If*

$$\lambda_2 \geq \frac{\delta}{\lambda}$$

*then the spectral gap of the transformed network structure is 0.*

This is interesting because the spectral gap has an interpretation in terms of network structure. A strictly positive spectral gap corresponds to a network with only one component. A small but positive spectral gap corresponds to a network with at least two communities with many within-group links and few between-group links. After intervention and under the condition stated in proposition 6 the spectral gap will close to zero. The new network will be characterized by communities that do not share links across them. The

social planner isolate players into separate communities and then reduces the spread of the disease inside each community until elimination.

It is important to note that Theorem 3 also gives us the cost of the transformation, as a simple function of  $\lambda, \delta$ , and the eigenvalues of  $G$ . It means that given two network structures and an epidemic threshold, we directly have a number enabling us to rank those two networks in terms of how costly it would be for the social planner to intervene. It can also be interpreted as how close each network is from a structure that prevents an epidemic outbreak. In the next section we will use this information to do comparative statics on different networks.

## 5 Applications

We discuss limitations and benefits of our method, and illustrate them in three examples hereafter.

When applying the intervention described in the previous section, we see that the post-intervention matrix described in equation (4.4) may have negative entries. Our method requires to make the matrix  $A(G) = G - \frac{\delta}{\lambda}$  negative semi-definite. There is not a condition that preclude the new matrix (after intervention) to exhibit negative entries. In the epidemics framework, a link with negative intensity between two individuals is translated into a decreased probability to be infected for one of the individuals when the other becomes infected. This interpretation is difficult to justify in some of the applications we described.

When the minimal-cost intervention we presented make some entries of the modified network negative, our theorem does not give information about other interventions to prevent epidemics diffusion. The theorem 2.1 of Higham (1988) [?] only addresses minimization problems that do not apply constraints on the outcome matrix properties. We therefore regard our result as a lower bound of the cost of intervention. The problem that the planner has to solve is the same one but with the additional constraint of having the post-intervention matrix with non-negative entries only. When this additional constraint is not binding, meaning that the matrix  $\tilde{G}$  defined in (4.4) has no negative entries, theorem 3 hits its full potential. It tells us what is the intervention that corresponds to the minimal-cost intervention, and gives us a simple formula of the cost of the intervention, allowing for quick comparative statics. We give hereafter examples where this lower bound for the cost of intervention is reached.



## 5.1 Example 1

### 5.1.1 Initial structure

We consider a population of  $n$  individuals who differ in their total intensity of interaction (or equivalently their probability of spreading the disease). Each individual  $i$  is characterized by a coefficient  $c_i$  measuring his propensity of interaction.  $c_i$  is drawn from a distribution  $\mathcal{C}$  with support  $\mathcal{R}^+$ . We assume random mixing, that is the probability of having an interaction with an individual  $j$  is proportional to  $j$ 's propensity of interaction. The strength of the link between individuals  $i$  and  $j$  is therefore:

$$g_{ij} = c_i c_j$$

Two individuals are very likely to have an interaction if they both have an high propensity of interaction. A different interpretation is that if  $c_i$  measures the contagiousness of individual  $i$  it is likely that an interaction with him will result in an high probability of infection. The network framework described in this example is the intensity counterpart of the degree distribution framework where links are non-weighted, 0 or 1, but exist with a probability depending on the degree of two nodes. If nodes  $i$  and  $j$  have degree  $d_i$  and  $d_j$  respectively, the probability that they are linked is proportional to  $d_i d_j$ . In our example, all nodes are linked, and the weight of the link between  $i$  and  $j$  is  $c_i c_j$ .

We will try to determine the optimal intervention of the planner and to compare the cost of intervention for different distributions. In order to that, we will compute the eigenvalues and eigenvectors of the adjacency matrix to apply Theorem 3.

**Eigenvalues and eigenvectors of  $G$**  The only positive eigenvalue of  $G$  is  $\lambda_1$ . For  $i = 2, \dots, n$ ,  $\lambda_i = 0$ . The eigenvector associated to  $\lambda_1$  is  $v_1$ . We have:

$$\lambda_1 = c_1^2 + c_2^2 + \dots + c_n^2$$

and

$$v_1 = \begin{bmatrix} \frac{c_1}{\sqrt{c_1^2 + c_2^2 + \dots + c_n^2}} \\ \frac{c_2}{\sqrt{c_1^2 + c_2^2 + \dots + c_n^2}} \\ \dots \\ \frac{c_n}{\sqrt{c_1^2 + c_2^2 + \dots + c_n^2}} \end{bmatrix}$$

### 5.1.2 Post-intervention structure

Applying theorem 3 we can see that in order to prevent an epidemics, the planner transforms the structure from  $G$  to  $\tilde{G}$ :

$$\tilde{G} = \frac{\delta}{c_1^2 + c_2^2 + \dots + c_n^2} G$$

This means that each link is affected proportionally to its initial intensity.

### 5.1.3 Comparative statics of the cost of intervention

We wish to derive comparative statics with respect to the distribution  $\mathcal{C}$  which determines the intensities  $\{c_i\}_i$ .

We are interested in comparative statics with respect to the variance of the  $\{c_i\}_i$ , as it is trivial that increasing the average of  $\mathcal{C}$  increases the diffusion of the disease and thus the cost of the intervention. The idea is to compare a network where individuals are similar in their propensity to interaction to one where individuals are instead heterogeneous.

In order to neutralize the effect of the mean, we compare mean-preserving spreads of  $\mathcal{C}$ . Given the weight of each  $c_i$  is one, the mean-preserving spread will express itself through the support of the  $\{c_i\}_i$ .

Theorem 3 gives us the cost of the intervention:

$$\|\tilde{G} - G\|_F^2 = \lambda_1 - \frac{\delta}{\lambda} = c_1^2 + c_2^2 + \dots + c_n^2$$

We see that by taking a mean-preserving spread of the initial distribution of  $\{c_i\}_i$ , we increase the cost of intervention.

The less spread the distribution  $\mathcal{C}$  is, the easier it is for the planner to intervene on the network. Another

interpretation is that as links intensities are more homogeneous, it is unambiguously easier to intervene to prevent epidemics.

**Comparison with Jackson, Rogers (2007) [?]** We wish to draw a parallel with proposition 2 of Jackson, Rogers (2007) [?] (hereafter JR), which states that when the epidemic characteristics  $\frac{\lambda}{\delta}$  is low enough, a mean-preserving spread of the degree distribution of the network yields higher average infection. This result echoes ours that by spreading the intensities, it becomes more difficult to immunize the population (in terms of higher cost of intervention). The comparison is delicate however because of the following two reasons:

- Both JR and we make comparative statics with respect to spreads of the degree distribution of the network. The object of interest is different though: they show results on the average level of infection in the (positive) steady-state, while we provide the cost of reaching a zero steady state level from an initial network that exhibits a positive level of infection. A parallel may thus be drawn between lowering the steady-state level of infection (in JR) and lowering the cost of reaching a zero-steady state (in this paper). Even though it intuitively makes sense, we don't have the proof of it.
- Proposition 2 of Jackson, Rogers (2007) [?] has two elements. It states the existence of two thresholds  $\underline{\lambda}$  and  $\bar{\lambda}$ . When  $\frac{\lambda}{\delta}$  is low enough ( $\frac{\lambda}{\delta} < \underline{\lambda}$ ), a spread in the degree distribution yields *higher average infection*. But when  $\frac{\lambda}{\delta}$  is high enough ( $\frac{\lambda}{\delta} > \bar{\lambda}$ ), a spread in the degree distribution yields *lower average infection*. In the first case, the behavior of JR's model and ours is comparable (comparable in the sense defined in the previous element). Taking a mean-preserving spread of the degree distribution increases the average level of infection in JR, and increases our cost of intervention. This statement naturally rises (raises?) the question of why we focus on the first case. Our answer is the following: it makes sense to think that what matters for total eradication is the case where  $\frac{\lambda}{\delta}$  is low enough and we have a low average infection, that is the first case of JR. However this claim disregards the fact that we don't know the dynamics of the average infection rate if we were to progressively increase (or decrease) the spread of the degree distribution. This matters if our initial starting point places us in the case where  $\frac{\lambda}{\delta} > \bar{\lambda}$ .

**Hypothetical dynamics when increasing the spread in Jackson, Rogers (2007) [?]** We start from  $\frac{\lambda}{\delta} > \bar{\lambda}$ . Figure 3 is a picture of the  $\frac{\lambda}{\delta}$  with respect to the thresholds:

In this case of JR, an increase in the spread of the degree distribution decreases average infection. To understand why a starting point at  $\frac{\lambda}{\delta} > \bar{\lambda}$  is delicate, we refer to the proof of proposition 2 of JR. They

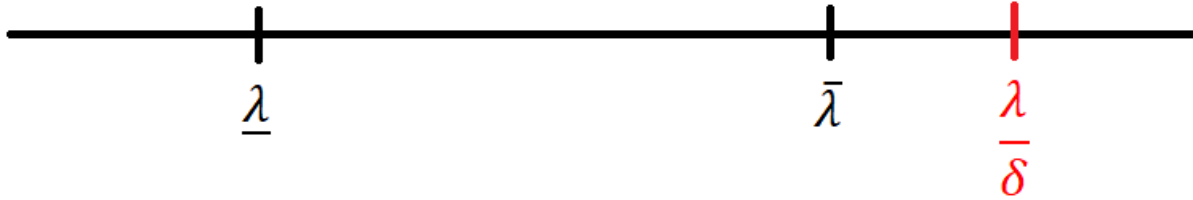


Figure 3: Thresholds for the cases in proposition 2, Jackson, Rogers (2007)

resort to an intermediary variable, that they call  $\theta$ , which is the average level of infection taken with respect to a transformation of the degree distribution. The existence of two opposite reactions to a mean-preserving spread comes from the fact that this variable  $\theta$  is always increasing with a mean-preserving spread, but the relationship between the actual average infection (that they call  $\rho$ ) and  $\theta$  is hump-shaped (see figure 4). The road towards the zero steady state from  $\theta = \theta_1$  (on the same figure 4) consists of continuously decreasing the mean-preserving spread, and thus  $\theta$ , even though the average steady-state  $\rho$  (for instance,  $\rho_2 > \rho_1$ ) increases first before decreasing. It is especially difficult to understand in which case one stands as when passing from  $\theta = \theta_2$  to  $\theta = \theta_3$ , one switches from the high zone to the low zone, but does not realize by observing the relationship between  $\theta$  and  $\rho$  (note that  $\rho_3 > \rho_2$  even though we are now in the zone where  $\rho$  and  $\theta$  co-move). Therefore we do not share the ambiguity of the role of the mean-preserving spread present in JR paper.

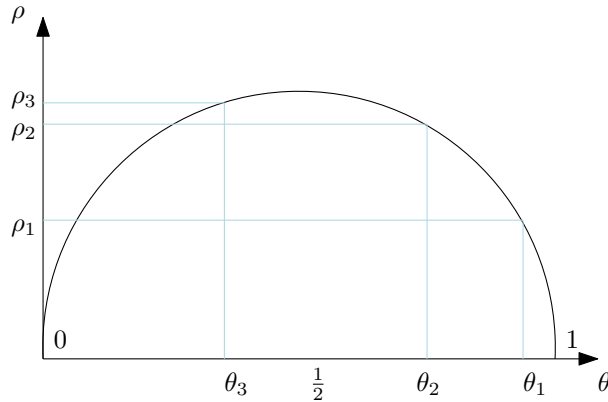


Figure 4: Relationship between average infection  $\rho$  and  $\theta$  while changing the degree distribution, at fixed  $\lambda$ , in Jackson, Rogers (2007)

**Insights from the comparison** This potentially cyclical relationship between spread and average infection limits the comparison between Jackson, Rogers' result and ours. Yet we refer their result as it

shed lights on one possible mechanism for our finding. In their paper, they claim that a spread in degree distribution boosts average infection because when the exogenous contagiousness is not favorable to the disease (low  $\frac{\lambda}{\delta}$ ), it is crucial to have very high degree nodes that serve as conductors of the disease, otherwise the epidemics would die out. If this behavior is the last resort of the disease before vanishing, it makes sense that the planner prevents it by containing the structural inequality in the network structure.

## 5.2 Example 2

### 5.2.1 Initial structure

We consider a population of  $n$  agents divided into two groups of equal size: agents  $1 \dots \frac{n}{2}$  belong to group 1 and agents  $\frac{n}{2} + 1, \dots, n$  belong to group 2. Intragroup links have strength  $1 - \epsilon$ , intergroup links have strength  $\epsilon$ , with  $\epsilon \leq 1/2$ . Therefore links are more intense within than between groups. Such a network is represented by the following adjacency matrix:

$$G = \begin{bmatrix} (1-\epsilon) & (1-\epsilon) & \dots & (1-\epsilon) & \epsilon & \dots & \epsilon & \epsilon \\ (1-\epsilon) & (1-\epsilon) & \dots & (1-\epsilon) & \epsilon & \dots & \epsilon & \epsilon \\ & & \ddots & & & & & \\ (1-\epsilon) & (1-\epsilon) & \dots & (1-\epsilon) & \epsilon & \dots & \epsilon & \epsilon \\ \epsilon & \epsilon & \dots & \epsilon & (1-\epsilon) & \dots & (1-\epsilon) & (1-\epsilon) \\ & & & & & \ddots & & \\ \epsilon & \epsilon & \dots & \epsilon & (1-\epsilon) & \dots & (1-\epsilon) & (1-\epsilon) \end{bmatrix} \quad (5.1)$$

$\epsilon$  represents the strength of inequality between the two types of links: intragroup and intergroup links. As  $\epsilon$  grows from 0 to  $\frac{1}{2}$ , the total intensity of links does not move, but the distribution of this intensity does, shifting from maximum heterogeneity (1 versus 0) to total homogeneity ( $\frac{1}{2}$  for all links). Note that by restricting  $\epsilon$  to be less or equal to  $\frac{1}{2}$ , we focus on positive assortative matching, and do not consider negative assortative matching (where individuals are more linked with members of the other group than with members of their own).

**Eigenvalues and eigenvectors of  $G$**  The matrix  $G$  has two positive eigenvalues:

$$\lambda_1 = \frac{n}{2}, \quad \lambda_2 = (1 - 2\epsilon)\frac{n}{2}, \quad \lambda_3 = \dots \lambda_n = 0$$

The associate eigenvectors are:

$$v_1 = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \dots \\ 1 \\ 1 \\ \dots \\ 1 \end{bmatrix}, \quad v_2 = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \dots \\ 1 \\ -1 \\ \dots \\ -1 \end{bmatrix}$$

The sign of the elements in  $v_2$  depends on the group of each individual. The  $\frac{n}{2}$  elements corresponding to individuals of the first group are positive, the  $\frac{n}{2}$  elements corresponding to individuals of the second groups are negative.

### 5.2.2 Post-intervention structure

The post-intervention structure  $\tilde{G}$  is:

$$\tilde{G} = \begin{bmatrix} \frac{\delta}{\lambda} \frac{2}{n} & \dots & \frac{\delta}{\lambda} \frac{2}{n} & 0 & \dots & 0 \\ \dots & & \dots & & & \\ \frac{\delta}{\lambda} \frac{2}{n} & \dots & \frac{\delta}{\lambda} \frac{2}{n} & 0 & \dots & 0 \\ 0 & \dots & 0 & \frac{\delta}{\lambda} \frac{2}{n} & \dots & \frac{\delta}{\lambda} \frac{2}{n} \\ & & & \dots & & \dots \\ 0 & \dots & 0 & \frac{\delta}{\lambda} \frac{2}{n} & \dots & \frac{\delta}{\lambda} \frac{2}{n} \end{bmatrix}$$

The first thing to note is that the outcome structure does not depend on  $\epsilon$ . Regardless of the value of  $\epsilon$ , the planner removes all the intergroup links, and set all the intragroup links to  $\frac{\delta}{\lambda} \frac{2}{n}$ .

### 5.2.3 Comparative statics of the cost of intervention

We want to study the behavior of the cost of intervention as a function of the inequality of relationships between intra and intergroup,  $\epsilon$ . Theorem 3 gives us the exact value of the cost  $K$  of the intervention:

$$K = \sum_{\lambda_r > \frac{\delta}{\lambda}} \left( \lambda_r - \frac{\delta}{\lambda} \right)^2 = \left( \frac{n}{2} - \frac{\delta}{\lambda} \right)^2 \mathbf{1}_{\{\frac{n}{2} > \frac{\delta}{\lambda}\}} + \left( (1-2\epsilon)\frac{n}{2} - \frac{\delta}{\lambda} \right)^2 \mathbf{1}_{\{(1-2\epsilon)\frac{n}{2} > \frac{\delta}{\lambda}\}}$$

We can rewrite the cost as a function of  $n$ :

$$K = \begin{cases} 0 & \text{if } n < 2\frac{\delta}{\lambda} & \text{(no eigenvalue is lowered)} \\ \left( \frac{n}{2} - \frac{\delta}{\lambda} \right)^2 & \text{if } 2\frac{\delta}{\lambda} \leq n < 2\frac{\delta}{\lambda} \frac{1}{1-2\epsilon} & \text{(one eigenvalue is lowered)} \\ \left( \frac{n}{2} - \frac{\delta}{\lambda} \right)^2 + \left( (1-2\epsilon)\frac{n}{2} - \frac{\delta}{\lambda} \right)^2 & \text{if } 2\frac{\delta}{\lambda} \frac{1}{1-2\epsilon} \leq n & \text{(two eigenvalues are lowered)} \end{cases} \quad (5.2)$$

Let us focus on the case where there is intervention (that is  $n > 2\frac{\delta}{\lambda}$ ). We call  $\bar{\epsilon}$  the following threshold for  $\epsilon$ :

$$\bar{\epsilon} = \frac{1}{2} - \frac{\delta}{\lambda n}$$

As a direct application of Theorem 3 we have:<sup>51</sup>

- For  $\epsilon \in [0, \bar{\epsilon}]$ , the cost of intervention strictly decreases with  $\epsilon$ , from  $\left( (1-2\epsilon)\frac{n}{2} - \frac{\delta}{\lambda} \right)^2$  to  $\left( \frac{n}{2} - \frac{\delta}{\lambda} \right)^2$
- For  $\epsilon > \bar{\epsilon}$ , the cost of intervention is constant as a function of  $\epsilon$ , at  $\left( \frac{n}{2} - \frac{\delta}{\lambda} \right)^2$

As a whole, the cost of intervention for the planner is weakly decreasing in  $\epsilon$  (strictly decreasing for  $\epsilon < \bar{\epsilon}$  and then stable). The more isolated the two communities are, the more difficult it is to eliminate the disease in this setup. One potential explanation would be that decreasing the diffusion in one group has spillovers on the other group, reducing the spread of the disease there too. The more linked the communities are ex-ante, the bigger this effect is. This can seem contradictory to the ex-post structure described in (5.4) where the intergroup links intensity is lowered from  $\epsilon$  to 0. Our result shows that the extra cost of intervention on intergroup links resulting from an increase in  $\epsilon$  is more than compensated by the extra saving made on intragroup links. We can see it analytically, by computing the derivative of the cost with respect to  $\epsilon$ . In order to separate the two effects, we rewrite the costs under another form, directly coming from the formula of the Frobenius norm for  $\|G - \tilde{G}\|_F^2$ . There is the same number of links of each type,  $\frac{n^2}{2}$ , we can thus compare the change of cost per link with respect to a change in  $\epsilon$ :

- Marginal cost of intragroup link change:  $2\epsilon$

<sup>51</sup>The decomposition of the support of  $\epsilon$  makes sense as  $\bar{\epsilon}$  is positive under assumption ???. For any  $\epsilon$  bigger than  $\bar{\epsilon}$ , the cost of intervention remains at  $\left( \frac{n}{2} - \frac{\delta}{\lambda} \right)^2$  as the second eigenvalue remains under the threshold  $\frac{\delta}{\lambda}$  and the intervention concentrates on lowering the highest eigenvalue  $\lambda_1$ , which is independent of  $\epsilon$ . However on  $[0, \bar{\epsilon}]$ , the second eigenvalue is beyond the threshold  $\frac{\delta}{\lambda}$ , decreases with  $\epsilon$  and consequently lowers the cost of intervention.

- Marginal cost of intergroup link change:  $2\epsilon - 2\left(1 - \frac{\delta}{\lambda} \frac{2}{n}\right)$

Therefore, the total marginal cost (divided by the number of links of each type) is:

$$4\epsilon - 2\left(1 - \frac{\delta}{\lambda} \frac{2}{n}\right) \quad (5.3)$$

which is strictly negative for  $\epsilon < \bar{\epsilon}$ . The cost savings on the intergroup links more than compensate for the extra cost on intragroups links when we increase  $\epsilon$ .

We can compare this effect of  $\epsilon$  on the cost of eradicating the disease with Galeotti, Rogers (2013) [?]. They intervene at the group level, as in this example. They consider an intervention where part of the individuals get vaccination and therefore cannot be infected, as opposed to our intervention that targets links. They find that, under positive assortative matching (our setup), the planner should spread its immunization effort equally across both groups (Proposition 2). Our result is compatible with theirs. The planner acts symmetrically with regard to groups. However their cost of immunization necessary to eradicate the disease does not depend on the relative weight of intragroup and intergroup links. It would be interesting to further study the difference between the two methods to understand whether passing from a  $n$  dimension intervention to a  $n^2$  dimension intervention grants substantial benefits.

### 5.3 Example 3: homogeneous versus heterogeneous networks

We want to generalize what we found in the previous two examples. In both cases, in fact, it results that for the planner it is more difficult to intervene when the population is heterogeneous in terms of the link structure.

We define homogeneous networks in the following way:

**Definition 8.** *Let  $G \in \mathcal{M}_{n,n}$  be the adjacency matrix of a network. If there exists a number  $x \in \mathbb{R}$ ,  $x \geq 0$  such that:*

$$g_{ij} = x \quad \text{for all } i, j$$

*then the network is said to be homogeneous.*

The network defined in Definition 2 is one where all the individuals are equally connected to the others. We want to compare this network with one where the population is instead heterogeneous.



The next proposition tell us that it is always cheaper to intervene in an homogeneous network:

**Proposition 10.** *Consider two networks where the sum of the intensity of all connections is equal to 1. The first network is homogeneous, the second is not. The cost of intervention in the first network is strictly smaller than the cost in the second.*

*Proof.* We can associate to the adjacency matrix of any homogeneous network a corresponding stochastic matrix and use the theory of Markov chains to derive some insights on the planner's intervention. (see Levin et al., 2006, [?])

When the sum of elements of each column of the matrix is constant and equal to 1, the largest eigenvalue is equal to 1 and the corresponding eigenvector is  $(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})$ .

The corresponding adjacency matrix is given by:

$$G = \lambda_1^* \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \dots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \dots & \frac{1}{\sqrt{n}} \\ & & \ddots & \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \dots & \frac{1}{\sqrt{n}} \end{bmatrix} \quad (5.4)$$

where  $\lambda_1^*$  si the eigenvalue of  $G$ . It can be rewritten as:

$$\lambda_1^* \frac{1}{\sqrt{n}} (\frac{1}{\sqrt{n}})^T$$

The adjacency matrix of a generic network can be written, using the singular value decomposition, as:

$$\lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T + \dots + \lambda_n v_n v_n^T$$

where  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the eigenvalues of the adjacency matrix and  $v_1, v_2, \dots, v_n$  are the corresponding eigenvectors.

We can compute the total number of connections in a network summing all the entries of the adjacency matrix. The number of links for the homogeneous network is:

$$\lambda_1^* (\frac{1}{\sqrt{n}})^T (\frac{1}{\sqrt{n}})$$

Similarly, summing all the entries of the matrices obtained from the eigendecomposition (weighted by the corresponding eigenvalues) we obtained that the number of links for the non-homogeneous network is:

$$\lambda_1(v_1^T(\frac{1}{\sqrt{n}}))^2 + \lambda_2(v_2^T(\frac{1}{\sqrt{n}}))^2 + \dots + \lambda_n(v_n^T(\frac{1}{\sqrt{n}}))^2$$

Given that we want to compare two network with the same number of links we ask that the two previous expressions are the same:

$$\lambda_1^* = \lambda_1^*(\frac{1}{\sqrt{n}})^T(\frac{1}{\sqrt{n}}) = \lambda_1(v_1^T(\frac{1}{\sqrt{n}}))^2 + \lambda_2(v_2^T(\frac{1}{\sqrt{n}}))^2 + \dots + \lambda_n(v_n^T(\frac{1}{\sqrt{n}}))^2$$

We note that the dot product  $v_i^T(\frac{1}{\sqrt{n}})$  can be written as  $\|v_i\| * \|\frac{1}{\sqrt{n}}\| \cos \theta_i$  where  $\theta_i$  is the angle between the vector  $\frac{1}{\sqrt{n}}$  and eigenvector  $v_i$ . Hence, we can express the RHS of the previous equation as:

$$\begin{aligned} \lambda_1 \|v_1\|^2 * \|\frac{1}{\sqrt{n}}\|^2 \cos^2 \theta_1 + \lambda_2 \|v_2\|^2 * \|\frac{1}{\sqrt{n}}\|^2 \cos^2 \theta_2 + \dots + \lambda_n \|v_n\|^2 * \|\frac{1}{\sqrt{n}}\|^2 \cos^2 \theta_n = \\ = \lambda_1 \cos^2 \theta_1 + \lambda_2 \cos^2 \theta_2 + \dots + \lambda_n \cos^2 \theta_n \end{aligned}$$

where the last equality comes from the fact that the eigenvectors have length 1 (from the singular value decomposition). Note also that  $\cos^2 \theta_i \in [0,1]$ . Finally observe that  $\cos^2 \theta_1 + \cos^2 \theta_2 + \dots + \cos^2 \theta_n = 1$ , being  $v_1, v_2, \dots, v_n$  an orthonormal basis for  $\mathbb{R}^n$

Therefore we know that  $\lambda_1^*$  can be written as a convex combination of the eigenvalues of the non-homogeneous network. This implies that at least one of the eigenvalue of the non-homogeneous network is larger than  $\lambda_1^*$ . As a direct consequence of Theorem 1, we know that the cost of intervention for the non-homogeneous network must be bigger.  $\square$

## 6 Conclusion

We investigate contagion processes among a networked population. We use results from linear algebra applied to computer science to find how to prevent contagion from an initial seed. For a given diffusion process and an arbitrary given initial network, we give the closest network to the initial one such that no outbreak occurs. We provide intuition on what the intervention on the network structure looks like by

analyzing relevant examples of connection patterns. We discuss the limitations of our result together with the potential for future research.